```
=> fil hcaplus
FILE 'HCAPLUS' ENTERED AT 08:53:06 ON 26 FEB 2003
USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.
PLEASE SEE "HELP USAGETERMS" FOR DETAILS.
COPYRIGHT (C) 2003 AMERICAN CHEMICAL SOCIETY (ACS)
```



```
FILE COVERS 1907 - 26 Feb 2003  VOL 138 ISS 9
FILE LAST UPDATED: 25 Feb 2003  (20030225/ED)

 This file contains CAS Registry Numbers for easy and accurate
 substance identification.


=>
=>

=> d stat que l7
L2           59 SEA FILE=REGISTRY ABB=ON  PLU=ON  SGTVGR/SQSP
L3           16 SEA FILE=REGISTRY ABB=ON  PLU=ON  SKLMYD/SQSP
L4           25 SEA FILE=REGISTRY ABB=ON  PLU=ON  L2 AND SQL=<30
L5            3 SEA FILE=HCAPLUS ABB=ON   PLU=ON  L3
L6            3 SEA FILE=HCAPLUS ABB=ON   PLU=ON  L4
L7            3 SEA FILE=HCAPLUS ABB=ON   PLU=ON  L5 OR L6

=> d ibib abs hitrn l7 1-3

L7   ANSWER 1 OF 3  HCAPLUS  COPYRIGHT 2003 ACS
ACCESSION NUMBER:         2000:898363  HCAPLUS
DOCUMENT NUMBER:          134:174115
TITLE:                    Cardiovascular and autonomic effects of
                          .omega.-conotoxins MVIIA and CVID in conscious rabbits
                          and isolated tissue assays
AUTHOR(S):                Wright, Christine E.; Robertson, Alan D.; Whorlow,
                          Sarah L.; Angus, James A.
CORPORATE SOURCE:         Department of Pharmacology, The University of
                          Melbourne, Victoria, 3010, Australia
SOURCE:                   British Journal of Pharmacology (2000), 131(7),
                          1325-1336
                          CODEN: BJPCBM; ISSN: 0007-1188
PUBLISHER:                Nature Publishing Group
DOCUMENT TYPE:            Journal
LANGUAGE:                 English
AB   The effects of a novel N-type voltage-operated calcium channel antagonist,
     .omega.-conotoxin CVID, were compared with .omega.-conotoxin MVIIA on
     sympathetic-evoked activation of right atria (RA), small mesenteric
     arteries (MA) and vasa deferentia (VD) isolated from the rat.  Their
     effects were also compared on blood pressure and cardiovascular reflexes
     in conscious rabbits.  The pIC50 values for MVIIA and CVID, resp., for
     inhibiting sympathetic-evoked responses were equiv. in RA (8.7 and 8.7)
     and VD (9.0 and 8.7); however, in MA the values were 8.4 and 7.7.  The
     cardiac to vascular (RA/MA) potency ratios, antilog (plog RA-plog MA), for
```

MVIIA and CVID were 2 and 10. The offset rates for CVID and MVIIA were rapid, and peptide reapplication caused rapid onset of blockade, suggesting limited desensitization. In the conscious rabbit, CVID and MVIIA (100 .mu.g kg-1 i.v.) caused a similar fall in blood pressure and a tachycardia that rapidly reached max. Both peptides decreased the vagal- and sympathetic-mediated components of the baroreflex, but had no effect on the vagal nasopharyngeal reflex. The orthostatic reflex to 90.degree. tilt was blocked by MVIIA with sustained postural hypotension for .gtoreq.90 min after administration. In contrast, CVID caused postural hypotension at 30 min which recovered rapidly. Neither CVID nor MVIIA (3 .mu.g kg-1 i.t.) significantly altered cardiovascular variables or autonomic reflexes. In conclusion, CVID appears to be relatively weak at inhibiting the reflex response to tilt consistent with its weaker inhibition of rat mesenteric artery constriction to perivascular nerve stimulation. This may point to subtype N-type calcium channel selectivity.

IT **247207-83-6**
RL: ADV (Adverse effect, including toxicity); BIOL (Biological study) (cardiovascular and autonomic effects of .omega.-conotoxins MVIIA and CVID in conscious rabbits and isolated tissue assays)

REFERENCE COUNT: 31 THERE ARE 31 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L7 ANSWER 2 OF 3 HCAPLUS COPYRIGHT 2003 ACS

ACCESSION NUMBER: 2000:813172 HCAPLUS
DOCUMENT NUMBER: 134:158676
TITLE: Novel .omega.-conotoxins from Conus catus discriminate among neuronal calcium channel subtypes
AUTHOR(S): Lewis, Richard J.; Nielsen, Katherine J.; Craik, David J.; Loughnan, Marion L.; Adams, Denise A.; Sharpe, Iain A.; Luchian, Tudor; Adams, David J.; Bond, Trudy; Thomas, Linda; Jones, Alun; Matheson, Jodi-Lea; Drinkwater, Roger; Andrews, Peter R.; Alewood, Paul F.
CORPORATE SOURCE: Centre for Drug Design and Development (3D Centre), Institute for Molecular Bioscience, Department of Physiology and Pharmacology, Queensland Agricultural Biotechnology Centre (QDPI), The University of Queensland, Brisbane, 4072, Australia
SOURCE: Journal of Biological Chemistry (2000), 275(45), 35335-35344
CODEN: JBCHA3; ISSN: 0021-9258
PUBLISHER: American Society for Biochemistry and Molecular Biology
DOCUMENT TYPE: Journal
LANGUAGE: English

AB .omega.-Conotoxins selective for N-type calcium channels are useful in the management of severe pain. In an attempt to expand the therapeutic potential of this class, four new .omega.-conotoxins (CVIA-D) have been discovered in the venom of the piscivorous cone snail, Conus catus, using assay-guided fractionation and gene cloning. Compared with other .omega.-conotoxins, CVID has a novel loop 4 sequence and the highest selectivity for N-type over P/Q-type calcium channels in radioligand binding assays. CVIA-D also inhibited contractions of elec. stimulated rat vas deferens. In electrophysiol. studies, .omega.-conotoxins CVID and MVIIA had similar potencies to inhibit current through central (.alpha.1B-d) and peripheral (.alpha.1B-b) splice variants of the rat N-type calcium channels when coexpressed with rat .beta.3 in Xenopus oocytes. However, the potency of CVID and MVIIA increased when .alpha.1B-d and .alpha.1B-b were expressed in the absence of rat .beta.3, an effect most pronounced for CVID at .alpha.1B-d (up to 540-fold) and least pronounced for MVIIA at .alpha.1B-d (3-fold). The novel selectivity of CVID may have therapeutic implications. 1H NMR studies reveal that CVID possesses a combination of unique structural features, including two

hydrogen bonds that stabilize loop 2 and place loop 2 proximal to loop 4, creating a globular surface that is rigid and well defined.

IT **325164-11-2P**, .omega.-Conotoxin C VIC
RL: ADV (Adverse effect, including toxicity); PRP (Properties); PUR (Purification or recovery); SPN (Synthetic preparation); BIOL (Biological study); PREP (Preparation)
(amino acid sequence; novel .omega.-conotoxins from Conus catus discriminate among neuronal calcium channel subtypes)

IT **247207-83-6P**, .omega.-Conotoxin C VID
RL: ADV (Adverse effect, including toxicity); PRP (Properties); PUR (Purification or recovery); SPN (Synthetic preparation); BIOL (Biological study); PREP (Preparation)
(amino acid sequence; .omega.-conotoxins from Conus catus discriminate among neuronal calcium channel subtypes)

REFERENCE COUNT: 59 THERE ARE 59 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L7 ANSWER 3 OF 3 HCAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 1999:691117 HCAPLUS
DOCUMENT NUMBER: 131:332121
TITLE: .omega.-Conotoxin peptides and use as therapeutic calcium channel blockers
INVENTOR(S): Drinkwater, Roger Desmond; Lewis, Richard James; Alewood, Paul Francis; Nielsen, Katherine Justine
PATENT ASSIGNEE(S): The University of Queensland, Australia
SOURCE: PCT Int. Appl., 81 pp.
CODEN: PIXXD2
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 1
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 9954350	A1	19991028	WO 1999-AU288	19990416
W: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM				
RW: GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG				
CA 2325803	AA	19991028	CA 1999-2325803	19990416
AU 9933227	A1	19991108	AU 1999-33227	19990416
AU 749617	B2	20020627		
EP 1071707	A1	20010131	EP 1999-914368	19990416
R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, FI				
JP 2002512254	T2	20020423	JP 2000-544688	19990416

PRIORITY APPLN. INFO.:
AU 1998-2989 A 19980416
AU 1999-8419 A 19990201
WO 1999-AU288 W 19990416

AB An isolated, synthetic or recombinant .omega.-conotoxin peptide is provided in which the fourth loop between cysteine residues 5 and 6 comprises SGTVGR or such a sequence which has undergone .gtoreq.1 amino acid substitutions or side chain modifications. The peptides of the invention are useful in the treatment of conditions benefiting from calcium channel blockade.

IT **249628-22-6P**
RL: BOC (Biological occurrence); BPN (Biosynthetic preparation); BSU (Biological study, unclassified); PRP (Properties); THU (Therapeutic use);

BIOL (Biological study); OCCU (Occurrence); PREP (Preparation); USES (Uses)
(amino acid sequence; .omega.-conotoxin peptides and use as therapeutic calcium channel blockers)

IT **247207-64-3P**
RL: BAC (Biological activity or effector, except adverse); BSU (Biological study, unclassified); PRP (Properties); PUR (Purification or recovery); SPN (Synthetic preparation); THU (Therapeutic use); BIOL (Biological study); PREP (Preparation); USES (Uses)
(.omega.-conotoxin peptides and use as therapeutic calcium channel blockers)

IT **247207-65-4P 247207-66-5P 247208-09-9P**
247208-10-2P 247208-11-3P
RL: BAC (Biological activity or effector, except adverse); BSU (Biological study, unclassified); PRP (Properties); SPN (Synthetic preparation); THU (Therapeutic use); BIOL (Biological study); PREP (Preparation); USES (Uses)
(.omega.-conotoxin peptides and use as therapeutic calcium channel blockers)

IT **247207-67-6 247207-68-7 247207-69-8**
247207-70-1 247207-71-2 247207-73-4
247207-75-6 247207-77-8 247207-79-0
247207-81-4 247207-83-6 247207-85-8
247207-87-0 247207-89-2 247207-91-6
247207-93-8 247207-95-0 247207-97-2
247207-99-4
RL: BAC (Biological activity or effector, except adverse); BSU (Biological study, unclassified); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); USES (Uses)
(.omega.-conotoxin peptides and use as therapeutic calcium channel blockers)

IT **247207-60-9 247207-61-0**
RL: BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study)
(.omega.-conotoxin peptides and use as therapeutic calcium channel blockers)

REFERENCE COUNT: 3 THERE ARE 3 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

=>
=>

=> fil reg
FILE 'REGISTRY' ENTERED AT 08:53:19 ON 26 FEB 2003

Property values tagged with IC are from the ZIC/VINITI data file provided by InfoChem.

STRUCTURE FILE UPDATES: 24 FEB 2003 HIGHEST RN 494745-03-8
DICTIONARY FILE UPDATES: 24 FEB 2003 HIGHEST RN 494745-03-8

TSCA INFORMATION NOW CURRENT THROUGH MAY 20, 2002

Please note that search-term pricing does apply when conducting SmartSELECT searches.

Crossover limits have been increased. See HELP CROSSOVER for details.

Experimental and calculated property data are now available. See HELP

PROPERTIES for more information. See STNote 27, Searching Properties in the CAS Registry File, for complete details:
http://www.cas.org/ONLINE/STN/STNOTES/stnotes27.pdf

=> s l3 or l4
L8 29 L3 OR L4

=> d .seq l8 1-29

L8 ANSWER 1 OF 29 REGISTRY COPYRIGHT 2003 ACS
RN 325164-11-2 REGISTRY
CN L-Cysteinamide, L-cysteinyl-L-lysylglycyl-L-lysylglycyl-L-glutaminyl-L-seryl-L-cysteinyl-L-seryl-L-lysyl-L-leucyl-L-methionyl-L-tyrosyl-L-.alpha.-aspartyl-L-cysteinyl-L-cysteinyl-L-threonylglycyl-L-seryl-L-cysteinyl-L-seryl-L-arginyl-L-arginylglycyl-L-lysyl-, cyclic (1.fwdarw.16),(8.fwdarw.20),(15.fwdarw.25)-tris(disulfide) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN .omega.-Conotoxin C VIC
CN .omega.-Conotoxin C VIC (Conus catus venom)
NTE modified

type	------ location ------		description
terminal mod.	Cys-26	-	C-terminal amide
bridge	Cys-1	- Cys-16	disulfide bridge
bridge	Cys-8	- Cys-20	disulfide bridge
bridge	Cys-15	- Cys-26	disulfide bridge

SQL 26

SEQ 1 CKGKGQSCSK LMYDCCTGSC SRRGKC
== ====
HITS AT: 9-14

REFERENCE 1: 134:158676

L8 ANSWER 2 OF 29 REGISTRY COPYRIGHT 2003 ACS
RN 249628-22-6 REGISTRY
CN .omega.-Conotoxin (Conus catus venom gene CVID) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN PN: WO9954350 FIGURE: 1 claimed protein
SQL 73

SEQ 51 AKCSKLMYDC CSGSCSGTVG RCG
======
HITS AT: 54-59

REFERENCE 1: 131:332121

L8 ANSWER 3 OF 29 REGISTRY COPYRIGHT 2003 ACS
RN 247208-11-3 REGISTRY
CN L-Cysteinamide, L-cysteinyl-L-lysyl-L-seryl-L-lysylglycyl-L-alanyl-L-lysyl-L-cysteinyl-L-.alpha.-aspartyl-L-arginyl-L-leucyl-L-methionyl-L-tyrosyl-L-.alpha.-aspartyl-L-cysteinyl-L-cysteinyl-L-serylglycyl-L-seryl-L-cysteinyl-L-serylglycyl-L-threonyl-L-valylglycyl-L-arginyl- (9CI) (CA INDEX NAME)
NTE modified

type	------ location ------		description
terminal mod.	Cys-27	-	C-terminal amide

SQL 27

SQL **27**

SEQ 1 CKSKGAKCDR LMYDCCSGSC SGTVGRC
======

HITS AT: 21-26

RELATED SEQUENCES AVAILABLE WITH SEQLINK

REFERENCE 1: 131:332121

L8 ANSWER 4 OF 29 REGISTRY COPYRIGHT 2003 ACS
RN 247208-10-2 REGISTRY
CN L-Cysteinamide, L-cysteinyl-L-lysyl-L-seryl-L-lysylglycyl-L-alanyl-L-lysyl-L-cysteinyl-L-seryl-L-arginyl-L-leucyl-L-methionyl-L-tyrosyl-L-.alpha.-aspartyl-L-cysteinyl-L-cysteinyl-L-serylglycyl-L-seryl-L-cysteinyl-L-serylglycyl-L-threonyl-L-valylglycyl-L-arginyl- (9CI) (CA INDEX NAME)
NTE modified

type	------ location ------	description
terminal mod.	Cys-27 -	C-terminal amide

SQL **27**
SQL **27**

SEQ 1 CKSKGAKCSR LMYDCCSGSC SGTVGRC
======

HITS AT: 21-26

RELATED SEQUENCES AVAILABLE WITH SEQLINK

REFERENCE 1: 131:332121

L8 ANSWER 5 OF 29 REGISTRY COPYRIGHT 2003 ACS
RN 247208-09-9 REGISTRY
CN L-Cysteinamide, L-cysteinyl-L-lysyl-L-seryl-L-lysylglycyl-L-alanyl-L-lysyl-L-cysteinyl-L-seryl-L-lysyl-L-leucyl-L-methionyl-L-tyrosyl-L-.alpha.-aspartyl-L-cysteinyl-L-cysteinyl-L-serylglycyl-L-seryl-L-cysteinyl-L-serylglycyl-L-threonyl-L-valylglycyl-L-arginyl- (9CI) (CA INDEX NAME)
NTE modified

type	------ location ------	description
terminal mod.	Cys-27 -	C-terminal amide

SQL **27**
SQL **27**

SEQ 1 CKSKGAKCSK LMYDCCSGSC SGTVGRC
== ==== ======

HITS AT: 9-14, 21-26

RELATED SEQUENCES AVAILABLE WITH SEQLINK

REFERENCE 1: 131:332121

L8 ANSWER 6 OF 29 REGISTRY COPYRIGHT 2003 ACS
RN 247207-99-4 REGISTRY
CN L-Cysteinamide, L-cysteinyl-L-lysyl-L-seryl-L-lysylglycyl-L-alanyl-L-lysyl-L-cysteinyl-L-seryl-L-lysyl-L-leucyl-(2S)-2-amino-4-(methylsulfinyl)butanoyl-L-tyrosyl-L-.alpha.-aspartyl-L-cysteinyl-L-cysteinyl-L-serylglycyl-L-seryl-L-cysteinyl-L-serylglycyl-L-threonyl-L-valylglycyl-L-arginyl-, cyclic (1.fwdarw.16),(8.fwdarw.20),(15.fwdarw.27)-

tris(disulfide) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN PN: WO9954350 SEQID: 32 claimed sequence
NTE modified

type	------ location ------		description
terminal mod.	Cys-27	-	C-terminal amide
bridge	Cys-1	- Cys-16	disulfide bridge
bridge	Cys-8	- Cys-20	disulfide bridge
bridge	Cys-15	- Cys-27	disulfide bridge
modification	Met-12	-	oxygen<O>

SQL **27**
SQL **27**

SEQ 1 CKSKGAKCSK LMYDCCSGSC SGTVGRC
== ==== ======
HITS AT: 9-14, 21-26

RELATED SEQUENCES AVAILABLE WITH SEQLINK

REFERENCE 1: 131:332121

L8 ANSWER 7 OF 29 REGISTRY COPYRIGHT 2003 ACS
RN 247207-97-2 REGISTRY
CN L-Cysteinamide, L-cysteinyl-L-lysyl-L-seryl-L-lysylglycyl-L-alanyl-L-lysyl-L-cysteinyl-L-seryl-L-lysyl-L-leucyl-O-methyl-L-seryl-L-tyrosyl-L-.alpha.-aspartyl-L-cysteinyl-L-cysteinyl-L-serylglycyl-L-seryl-L-cysteinyl-L-serylglycyl-L-threonyl-L-valylglycyl-L-arginyl-, cyclic (1.fwdarw.16),(8.fwdarw.20),(15.fwdarw.27)-tris(disulfide) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN PN: WO9954350 SEQID: 31 claimed sequence
NTE modified

type	------ location ------		description
terminal mod.	Cys-27	-	C-terminal amide
bridge	Cys-1	- Cys-16	disulfide bridge
bridge	Cys-8	- Cys-20	disulfide bridge
bridge	Cys-15	- Cys-27	disulfide bridge
modification	Ser-12	-	methyl<Me>

SQL **27**
SQL **27**

SEQ 1 CKSKGAKCSK LSYDCCSGSC SGTVGRC
======
HITS AT: 21-26

REFERENCE 1: 131:332121

L8 ANSWER 8 OF 29 REGISTRY COPYRIGHT 2003 ACS
RN 247207-95-0 REGISTRY
CN L-Cysteinamide, L-cysteinyl-L-lysyl-L-seryl-L-lysylglycyl-L-alanyl-L-lysyl-L-cysteinyl-L-seryl-L-lysyl-L-leucyl-O-methyl-L-homoseryl-L-tyrosyl-L-.alpha.-aspartyl-L-cysteinyl-L-cysteinyl-L-serylglycyl-L-seryl-L-cysteinyl-L-serylglycyl-L-threonyl-L-valylglycyl-L-arginyl-, cyclic (1.fwdarw.16),(8.fwdarw.20),(15.fwdarw.27)-tris(disulfide) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN PN: WO9954350 SEQID: 30 claimed sequence

NTE modified

type	location		description
terminal mod.	Cys-27	-	C-terminal amide
bridge	Cys-1	- Cys-16	disulfide bridge
bridge	Cys-8	- Cys-20	disulfide bridge
bridge	Cys-15	- Cys-27	disulfide bridge
uncommon	Hse-12	-	-
modification	Hse-12	-	methyl<Me>

SQL **27**
SQL **27**

SEQ 1 CKSKGAKCSK LXYDCCSGSC SGTVGRC
======

HITS AT: 21-26

REFERENCE 1: 131:332121

L8 ANSWER 9 OF 29 REGISTRY COPYRIGHT 2003 ACS
RN 247207-93-8 REGISTRY
CN L-Cysteinamide, L-cysteinyl-L-lysyl-L-tyrosyl-L-lysylglycyl-L-alanyl-L-lysyl-L-cysteinyl-L-seryl-L-arginyl-L-leucyl-L-norleucyl-L-tyrosyl-L-.alpha.-aspartyl-L-cysteinyl-L-cysteinyl-L-serylglycyl-L-seryl-L-cysteinyl-L-serylglycyl-L-threonyl-L-valylglycyl-L-arginyl-, cyclic (1.fwdarw.16),(8.fwdarw.20),(15.fwdarw.27)-tris(disulfide) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN PN: WO9954350 SEQID: 29 claimed sequence
NTE modified

type	location		description
terminal mod.	Cys-27	-	C-terminal amide
bridge	Cys-1	- Cys-16	disulfide bridge
bridge	Cys-8	- Cys-20	disulfide bridge
bridge	Cys-15	- Cys-27	disulfide bridge
uncommon	Nle-12	-	-

SQL **27**
SQL **27**

SEQ 1 CKYKGAKCSR LXYDCCSGSC SGTVGRC
======

HITS AT: 21-26

REFERENCE 1: 131:332121

L8 ANSWER 10 OF 29 REGISTRY COPYRIGHT 2003 ACS
RN 247207-91-6 REGISTRY
CN L-Cysteinamide, L-cysteinyl-L-lysyl-L-seryl-L-lysylglycyl-L-alanyl-L-lysyl-L-cysteinyl-L-seryl-L-arginyl-L-leucyl-L-norleucyl-L-tyrosyl-L-.alpha.-aspartyl-L-cysteinyl-L-cysteinyl-L-serylglycyl-L-seryl-L-cysteinyl-L-serylglycyl-L-threonyl-L-valylglycyl-L-arginyl-, cyclic (1.fwdarw.16),(8.fwdarw.20),(15.fwdarw.27)-tris(disulfide) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN PN: WO9954350 SEQID: 28 claimed sequence
NTE modified

type	location		description

terminal mod.	Cys-27	-	C-terminal amide
bridge	Cys-1	- Cys-16	disulfide bridge
bridge	Cys-8	- Cys-20	disulfide bridge
bridge	Cys-15	- Cys-27	disulfide bridge
uncommon	Nle-12	-	-

SQL **27**
SQL **27**

SEQ 1 CKSKGAKCSR LXYDCCSGSC SGTVGRC
======

HITS AT: 21-26

REFERENCE 1: 131:332121

L8 ANSWER 11 OF 29 REGISTRY COPYRIGHT 2003 ACS
RN 247207-89-2 REGISTRY
CN L-Cysteinamide, L-cysteinyl-L-lysyl-L-seryl-L-lysylglycyl-L-alanyl-L-lysyl-L-cysteinyl-L-seryl-L-lysyl-L-leucyl-L-norleucyl-L-tyrosyl-L-.alpha.-aspartyl-L-cysteinyl-L-cysteinyl-L-serylglycyl-L-seryl-L-cysteinyl-L-serylglycyl-L-threonyl-L-valylglycyl-L-arginyl-, cyclic (1.fwdarw.16),(8.fwdarw.20),(15.fwdarw.27)-tris(disulfide) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN PN: WO9954350 SEQID: 27 claimed sequence
NTE modified

type	------ location ------		description
terminal mod.	Cys-27	-	C-terminal amide
bridge	Cys-1	- Cys-16	disulfide bridge
bridge	Cys-8	- Cys-20	disulfide bridge
bridge	Cys-15	- Cys-27	disulfide bridge
uncommon	Nle-12	-	-

SQL **27**
SQL **27**

SEQ 1 CKSKGAKCSK LXYDCCSGSC SGTVGRC
======

HITS AT: 21-26

REFERENCE 1: 131:332121

L8 ANSWER 12 OF 29 REGISTRY COPYRIGHT 2003 ACS
RN 247207-87-0 REGISTRY
CN L-Cysteinamide, N-acetyl-L-cysteinyl-L-lysyl-L-seryl-L-lysylglycyl-L-alanyl-L-lysyl-L-cysteinyl-L-seryl-L-lysyl-L-leucyl-L-methionyl-L-tyrosyl-L-.alpha.-aspartyl-L-cysteinyl-L-cysteinyl-L-serylglycyl-L-seryl-L-cysteinyl-L-serylglycyl-L-threonyl-L-valylglycyl-L-arginyl-, cyclic (1.fwdarw.16),(8.fwdarw.20),(15.fwdarw.27)-tris(disulfide) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN PN: WO9954350 SEQID: 26 claimed sequence
NTE modified

type	------ location ------		description
terminal mod.	Cys-1	-	N-acetyl
terminal mod.	Cys-27	-	C-terminal amide
bridge	Cys-1	- Cys-16	disulfide bridge
bridge	Cys-8	- Cys-20	disulfide bridge
bridge	Cys-15	- Cys-27	disulfide bridge

SQL **27**
SQL **27**

```
SEQ      1 CKSKGAKCSK LMYDCCSGSC SGTVGRC
                   == ====       ======
HITS AT:   9-14, 21-26
```

RELATED SEQUENCES AVAILABLE WITH SEQLINK

REFERENCE 1: 131:332121

L8 ANSWER 13 OF 29 REGISTRY COPYRIGHT 2003 ACS
RN 247207-85-8 REGISTRY
CN L-Cysteinamide, L-tyrosyl-L-cysteinyl-L-lysyl-L-seryl-L-lysylglycyl-L-alanyl-L-lysyl-L-cysteinyl-L-seryl-L-lysyl-L-leucyl-L-methionyl-L-tyrosyl-L-.alpha.-aspartyl-L-cysteinyl-L-cysteinyl-L-serylglycyl-L-seryl-L-cysteinyl-L-serylglycyl-L-threonyl-L-valylglycyl-L-arginyl-, cyclic (2.fwdarw.17),(9.fwdarw.21),(16.fwdarw.28)-tris(disulfide) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN PN: WO9954350 SEQID: 25 claimed sequence
NTE modified

type	------ location ------		description
terminal mod.	Cys-28	-	C-terminal amide
bridge	Cys-2	- Cys-17	disulfide bridge
bridge	Cys-9	- Cys-21	disulfide bridge
bridge	Cys-16	- Cys-28	disulfide bridge

SQL **28**
SQL **28**

```
SEQ      1 YCKSKGAKCS KLMYDCCSGS CSGTVGRC
                 = =====       ======
HITS AT:   10-15, 22-27
```

REFERENCE 1: 131:332121

L8 ANSWER 14 OF 29 REGISTRY COPYRIGHT 2003 ACS
RN 247207-83-6 REGISTRY
CN L-Cysteine, L-cysteinyl-L-lysyl-L-seryl-L-lysylglycyl-L-alanyl-L-lysyl-L-cysteinyl-L-seryl-L-lysyl-L-leucyl-L-methionyl-L-tyrosyl-L-.alpha.-aspartyl-L-cysteinyl-L-cysteinyl-L-serylglycyl-L-seryl-L-cysteinyl-L-serylglycyl-L-threonyl-L-valylglycyl-L-arginyl-, cyclic (1.fwdarw.16),(8.fwdarw.20),(15.fwdarw.27)-tris(disulfide) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN .omega.-Conotoxin C VID
CN .omega.-Conotoxin C VID (Conus catus venom)
CN PN: WO9954350 SEQID: 24 claimed sequence
NTE

type	------ location ------		description
bridge	Cys-1	- Cys-16	disulfide bridge
bridge	Cys-8	- Cys-20	disulfide bridge
bridge	Cys-15	- Cys-27	disulfide bridge

SQL **27**
SQL **27**

```
SEQ      1 CKSKGAKCSK LMYDCCSGSC SGTVGRC
                   == ====       ======
HITS AT:   9-14, 21-26

**RELATED SEQUENCES AVAILABLE WITH SEQLINK**

REFERENCE   1:  134:174115

REFERENCE   2:  134:158676

REFERENCE   3:  131:332121

L8   ANSWER 15 OF 29  REGISTRY  COPYRIGHT 2003 ACS
RN   247207-81-4  REGISTRY
CN   L-Tyrosinamide, L-cysteinyl-L-lysyl-L-seryl-L-lysylglycyl-L-alanyl-L-lysyl-
     L-cysteinyl-L-seryl-L-lysyl-L-leucyl-L-methionyl-L-tyrosyl-L-.alpha.-
     aspartyl-L-cysteinyl-L-cysteinyl-L-serylglycyl-L-seryl-L-cysteinyl-L-
     serylglycyl-L-threonyl-L-valylglycyl-L-arginyl-L-cysteinyl-, cyclic
     (1.fwdarw.16),(8.fwdarw.20),(15.fwdarw.27)-tris(disulfide) (9CI)  (CA
     INDEX NAME)
OTHER NAMES:
CN   PN: WO9954350 SEQID: 23 claimed sequence
NTE  modified
```

type	------ location ------		description
terminal mod.	Tyr-28	-	C-terminal amide
bridge	Cys-1	- Cys-16	disulfide bridge
bridge	Cys-8	- Cys-20	disulfide bridge
bridge	Cys-15	- Cys-27	disulfide bridge

```
SQL  28
SQL  28

SEQ      1 CKSKGAKCSK LMYDCCSGSC SGTVGRCY
                   == ====       ======
HITS AT:   9-14, 21-26

REFERENCE   1:  131:332121

L8   ANSWER 16 OF 29  REGISTRY  COPYRIGHT 2003 ACS
RN   247207-79-0  REGISTRY
CN   L-Cysteinamide, L-cysteinyl-L-lysyl-L-seryl-L-lysyl-D-alanyl-L-alanyl-L-
     lysyl-L-cysteinyl-L-seryl-L-lysyl-L-leucyl-L-methionyl-L-tyrosyl-L-.alpha.-
     aspartyl-L-cysteinyl-L-cysteinyl-L-serylglycyl-L-seryl-L-cysteinyl-L-
     serylglycyl-L-threonyl-L-valylglycyl-L-arginyl-, cyclic
     (1.fwdarw.16),(8.fwdarw.20),(15.fwdarw.27)-tris(disulfide) (9CI)  (CA
     INDEX NAME)
OTHER NAMES:
CN   PN: WO9954350 SEQID: 22 claimed sequence
NTE  modified
```

type	------ location ------		description
terminal mod.	Cys-27	-	C-terminal amide
bridge	Cys-1	- Cys-16	disulfide bridge
bridge	Cys-8	- Cys-20	disulfide bridge
bridge	Cys-15	- Cys-27	disulfide bridge

```
SQL  27
SQL  27

SEQ      1 CKSKAAKCSK LMYDCCSGSC SGTVGRC
```

```
                  == ====       ======
HITS AT:   9-14, 21-26

REFERENCE   1:  131:332121

L8   ANSWER 17 OF 29  REGISTRY  COPYRIGHT 2003 ACS
RN   247207-77-8  REGISTRY
CN   L-Cysteinamide, L-cysteinyl-L-lysyl-L-seryl-L-lysylglycyl-L-alanyl-L-lysyl-
     L-cysteinyl-L-seryl-L-lysyl-L-leucyl-L-methionyl-L-tyrosyl-L-.alpha.-
     aspartyl-L-cysteinyl-L-cysteinyl-L-threonylglycyl-L-seryl-L-cysteinyl-L-
     serylglycyl-L-threonyl-L-valylglycyl-L-arginyl-, cyclic
     (1.fwdarw.16),(8.fwdarw.20),(15.fwdarw.27)-tris(disulfide) (9CI)  (CA
     INDEX NAME)
OTHER NAMES:
CN   PN: WO9954350 SEQID: 21 claimed sequence
NTE  modified
----------------------------------------------------------------------
 type            ------ location ------         description
----------------------------------------------------------------------
terminal mod.    Cys-27          -         C-terminal amide
bridge           Cys-1       - Cys-16      disulfide bridge
bridge           Cys-8       - Cys-20      disulfide bridge
bridge           Cys-15      - Cys-27      disulfide bridge
----------------------------------------------------------------------
SQL  27
SQL  27

SEQ      1 CKSKGAKCSK LMYDCCTGSC SGTVGRC
                  == ====       ======
HITS AT:   9-14, 21-26

REFERENCE   1:  131:332121

L8   ANSWER 18 OF 29  REGISTRY  COPYRIGHT 2003 ACS
RN   247207-75-6  REGISTRY
CN   L-Cysteinamide, L-cysteinyl-L-lysyl-L-seryl-L-lysylglycyl-L-alanyl-L-lysyl-
     L-cysteinyl-L-seryl-L-lysyl-L-leucyl-L-alanyl-L-tyrosyl-L-.alpha.-aspartyl-
     L-cysteinyl-L-cysteinyl-L-serylglycyl-L-seryl-L-cysteinyl-L-serylglycyl-L-
     threonyl-L-valylglycyl-L-arginyl-, cyclic (1.fwdarw.16),(8.fwdarw.20),(15.
     fwdarw.27)-tris(disulfide) (9CI)  (CA INDEX NAME)
OTHER NAMES:
CN   PN: WO9954350 SEQID: 20 claimed sequence
NTE  modified
----------------------------------------------------------------------
 type            ------ location ------         description
----------------------------------------------------------------------
terminal mod.    Cys-27          -         C-terminal amide
bridge           Cys-1       - Cys-16      disulfide bridge
bridge           Cys-8       - Cys-20      disulfide bridge
bridge           Cys-15      - Cys-27      disulfide bridge
----------------------------------------------------------------------
SQL  27
SQL  27

SEQ      1 CKSKGAKCSK LAYDCCSGSC SGTVGRC
                      ======
HITS AT:   21-26

REFERENCE   1:  131:332121

L8   ANSWER 19 OF 29  REGISTRY  COPYRIGHT 2003 ACS
RN   247207-73-4  REGISTRY
CN   L-Cysteinamide, L-cysteinyl-L-lysyl-L-tyrosyl-L-lysylglycyl-L-alanyl-L-
```

lysyl-L-cysteinyl-L-seryl-L-arginyl-L-leucyl-L-methionyl-L-tyrosyl-L-.alpha.-aspartyl-L-cysteinyl-L-cysteinyl-L-serylglycyl-L-seryl-L-cysteinyl-L-serylglycyl-L-threonyl-L-valylglycyl-L-arginyl-, cyclic (1.fwdarw.16),(8.fwdarw.20),(15.fwdarw.27)-tris(disulfide) (9CI) (CA INDEX NAME)

OTHER NAMES:

CN PN: WO9954350 SEQID: 19 claimed sequence

NTE modified

type	------ location ------		description
terminal mod.	Cys-27	-	C-terminal amide
bridge	Cys-1	- Cys-16	disulfide bridge
bridge	Cys-8	- Cys-20	disulfide bridge
bridge	Cys-15	- Cys-27	disulfide bridge

SQL **27**

SQL **27**

SEQ 1 CKYKGAKCSR LMYDCCSGSC SGTVGRC

======

HITS AT: 21-26

REFERENCE 1: 131:332121

L8 ANSWER 20 OF 29 REGISTRY COPYRIGHT 2003 ACS

RN 247207-71-2 REGISTRY

CN L-Cysteinamide, L-cysteinyl-L-lysyl-L-seryl-L-lysylglycyl-L-alanyl-L-lysyl-L-cysteinyl-L-.alpha.-aspartyl-L-lysyl-L-leucyl-L-methionyl-L-tyrosyl-L-.alpha.-aspartyl-L-cysteinyl-L-cysteinyl-L-serylglycyl-L-seryl-L-cysteinyl-L-serylglycyl-L-threonyl-L-valylglycyl-L-arginyl-, cyclic (1.fwdarw.16),(8.fwdarw.20),(15.fwdarw.27)-tris(disulfide) (9CI) (CA INDEX NAME)

OTHER NAMES:

CN PN: WO9954350 SEQID: 18 claimed sequence

NTE modified

type	------ location ------		description
terminal mod.	Cys-27	-	C-terminal amide
bridge	Cys-1	- Cys-16	disulfide bridge
bridge	Cys-8	- Cys-20	disulfide bridge
bridge	Cys-15	- Cys-27	disulfide bridge

SQL **27**

SQL **27**

SEQ 1 CKSKGAKCDK LMYDCCSGSC SGTVGRC

======

HITS AT: 21-26

REFERENCE 1: 131:332121

L8 ANSWER 21 OF 29 REGISTRY COPYRIGHT 2003 ACS

RN 247207-70-1 REGISTRY

CN L-Cysteinamide, L-cysteinyl-L-lysyl-L-seryl-L-lysylglycyl-L-alanyl-L-lysyl-L-cysteinyl-L-seryl-L-lysyl-L-leucyl-L-methionyl-L-tyrosyl-L-.alpha.-aspartyl-L-cysteinyl-L-cysteinyl-L-serylglycyl-L-seryl-L-cysteinyl-L-serylglycyl-L-alanyl-L-valylglycyl-L-arginyl-, cyclic (1.fwdarw.16),(8.fwdarw.20),(15.fwdarw.27)-tris(disulfide) (9CI) (CA INDEX NAME)

OTHER NAMES:

CN PN: WO9954350 SEQID: 17 claimed sequence

NTE modified

type	location		description
terminal mod.	Cys-27	-	C-terminal amide
bridge	Cys-1	- Cys-16	disulfide bridge
bridge	Cys-8	- Cys-20	disulfide bridge
bridge	Cys-15	- Cys-27	disulfide bridge

SQL 27

SEQ 1 CKSKGAKCSK LMYDCCSGSC SGAVGRC
== ====
HITS AT: 9-14

REFERENCE 1: 131:332121

L8 ANSWER 22 OF 29 REGISTRY COPYRIGHT 2003 ACS
RN 247207-69-8 REGISTRY
CN L-Cysteinamide, L-cysteinyl-L-lysyl-L-seryl-L-lysylglycyl-L-alanyl-L-glutaminyl-L-cysteinyl-L-seryl-L-lysyl-L-leucyl-L-methionyl-L-tyrosyl-L-.alpha.-aspartyl-L-cysteinyl-L-cysteinyl-L-serylglycyl-L-seryl-L-cysteinyl-L-serylglycyl-L-threonyl-L-valylglycyl-L-arginyl-, cyclic (1.fwdarw.16),(8.fwdarw.20),(15.fwdarw.27)-tris(disulfide) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN PN: WO9954350 SEQID: 16 claimed sequence
NTE modified

type	location		description
terminal mod.	Cys-27	-	C-terminal amide
bridge	Cys-1	- Cys-16	disulfide bridge
bridge	Cys-8	- Cys-20	disulfide bridge
bridge	Cys-15	- Cys-27	disulfide bridge

SQL **27**
SQL **27**

SEQ 1 CKSKGAQCSK LMYDCCSGSC SGTVGRC
== ==== ======
HITS AT: 9-14, 21-26

REFERENCE 1: 131:332121

L8 ANSWER 23 OF 29 REGISTRY COPYRIGHT 2003 ACS
RN 247207-68-7 REGISTRY
CN L-Cysteinamide, L-cysteinyl-L-lysyl-L-seryl-L-lysylglycyl-L-alanyl-L-arginyl-L-cysteinyl-L-seryl-L-lysyl-L-leucyl-L-methionyl-L-tyrosyl-L-.alpha.-aspartyl-L-cysteinyl-L-cysteinyl-L-serylglycyl-L-seryl-L-cysteinyl-L-serylglycyl-L-threonyl-L-valylglycyl-L-arginyl-, cyclic (1.fwdarw.16),(8.fwdarw.20),(15.fwdarw.27)-tris(disulfide) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN PN: WO9954350 SEQID: 15 claimed sequence
NTE modified

type	location		description
terminal mod.	Cys-27	-	C-terminal amide
bridge	Cys-1	- Cys-16	disulfide bridge
bridge	Cys-8	- Cys-20	disulfide bridge
bridge	Cys-15	- Cys-27	disulfide bridge

SQL **27**
SQL **27**

```
SEQ      1 CKSKGARCSK LMYDCCSGSC SGTVGRC
                   == ====       ======
HITS AT:   9-14, 21-26
```

REFERENCE 1: 131:332121

L8 ANSWER 24 OF 29 REGISTRY COPYRIGHT 2003 ACS
RN 247207-67-6 REGISTRY
CN L-Cysteinamide, L-cysteinyl-L-arginyl-L-seryl-L-lysylglycyl-L-alanyl-L-lysyl-L-cysteinyl-L-seryl-L-lysyl-L-leucyl-L-methionyl-L-tyrosyl-L-.alpha.-aspartyl-L-cysteinyl-L-cysteinyl-L-serylglycyl-L-seryl-L-cysteinyl-L-serylglycyl-L-threonyl-L-valylglycyl-L-arginyl-, cyclic (1.fwdarw.16),(8.fwdarw.20),(15.fwdarw.27)-tris(disulfide) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN PN: WO9954350 SEQID: 14 claimed sequence
NTE modified

type	------ location ------		description
terminal mod.	Cys-27	-	C-terminal amide
bridge	Cys-1	- Cys-16	disulfide bridge
bridge	Cys-8	- Cys-20	disulfide bridge
bridge	Cys-15	- Cys-27	disulfide bridge

SQL **27**
SQL **27**

```
SEQ      1 CRSKGAKCSK LMYDCCSGSC SGTVGRC
                   == ====       ======
HITS AT:   9-14, 21-26
```

REFERENCE 1: 131:332121

L8 ANSWER 25 OF 29 REGISTRY COPYRIGHT 2003 ACS
RN 247207-66-5 REGISTRY
CN L-Cysteinamide, L-cysteinyl-L-lysyl-L-seryl-L-lysylglycyl-L-alanyl-L-lysyl-L-cysteinyl-L-.alpha.-aspartyl-L-arginyl-L-leucyl-L-methionyl-L-tyrosyl-L-.alpha.-aspartyl-L-cysteinyl-L-cysteinyl-L-serylglycyl-L-seryl-L-cysteinyl-L-serylglycyl-L-threonyl-L-valylglycyl-L-arginyl-, cyclic (1.fwdarw.16),(8.fwdarw.20),(15.fwdarw.27)-tris(disulfide) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN PN: WO9954350 SEQID: 7 claimed sequence
NTE modified

type	------ location ------		description
terminal mod.	Cys-27	-	C-terminal amide
bridge	Cys-1	- Cys-16	disulfide bridge
bridge	Cys-8	- Cys-20	disulfide bridge
bridge	Cys-15	- Cys-27	disulfide bridge

SQL **27**
SQL **27**

```
SEQ      1 CKSKGAKCDR LMYDCCSGSC SGTVGRC
                              ======
HITS AT:   21-26
```

RELATED SEQUENCES AVAILABLE WITH SEQLINK

REFERENCE 1: 131:332121

L8 ANSWER 26 OF 29 REGISTRY COPYRIGHT 2003 ACS
RN 247207-65-4 REGISTRY
CN L-Cysteinamide, L-cysteinyl-L-lysyl-L-seryl-L-lysylglycyl-L-alanyl-L-lysyl-L-cysteinyl-L-seryl-L-arginyl-L-leucyl-L-methionyl-L-tyrosyl-L-.alpha.-aspartyl-L-cysteinyl-L-cysteinyl-L-serylglycyl-L-seryl-L-cysteinyl-L-serylglycyl-L-threonyl-L-valylglycyl-L-arginyl-, cyclic (1.fwdarw.16),(8.fwdarw.20),(15.fwdarw.27)-tris(disulfide) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN PN: WO9954350 SEQID: 6 claimed sequence
NTE modified

type	------ location ------		description
terminal mod.	Cys-27	-	C-terminal amide
bridge	Cys-1	- Cys-16	disulfide bridge
bridge	Cys-8	- Cys-20	disulfide bridge
bridge	Cys-15	- Cys-27	disulfide bridge

SQL **27**
SQL **27**

SEQ 1 CKSKGAKCSR LMYDCCSGSC SGTVGRC
======
HITS AT: 21-26

RELATED SEQUENCES AVAILABLE WITH SEQLINK

REFERENCE 1: 131:332121

L8 ANSWER 27 OF 29 REGISTRY COPYRIGHT 2003 ACS
RN 247207-64-3 REGISTRY
CN L-Cysteinamide, L-cysteinyl-L-lysyl-L-seryl-L-lysylglycyl-L-alanyl-L-lysyl-L-cysteinyl-L-seryl-L-lysyl-L-leucyl-L-methionyl-L-tyrosyl-L-.alpha.-aspartyl-L-cysteinyl-L-cysteinyl-L-serylglycyl-L-seryl-L-cysteinyl-L-serylglycyl-L-threonyl-L-valylglycyl-L-arginyl-, cyclic (1.fwdarw.16),(8.fwdarw.20),(15.fwdarw.27)-tris(disulfide) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN Leconotide
CN PN: WO9954350 SEQID: 5 claimed sequence
NTE modified

type	------ location ------		description
terminal mod.	Cys-27	-	C-terminal amide
bridge	Cys-1	- Cys-16	disulfide bridge
bridge	Cys-8	- Cys-20	disulfide bridge
bridge	Cys-15	- Cys-27	disulfide bridge

SQL **27**
SQL **27**

SEQ 1 CKSKGAKCSK LMYDCCSGSC SGTVGRC
== ==== ======
HITS AT: 9-14, 21-26

RELATED SEQUENCES AVAILABLE WITH SEQLINK

REFERENCE 1: 131:332121

L8 ANSWER 28 OF 29 REGISTRY COPYRIGHT 2003 ACS
RN 247207-61-0 REGISTRY
CN L-Aspartic acid, L-seryl-L-lysyl-L-leucyl-L-methionyl-L-tyrosyl- (9CI) (CA INDEX NAME)
OTHER NAMES:
CN PN: WO9954350 SEQID: 2 claimed sequence
SQL 6

SEQ 1 SKLMYD
======
HITS AT: 1-6

REFERENCE 1: 131:332121

L8 ANSWER 29 OF 29 REGISTRY COPYRIGHT 2003 ACS
RN 247207-60-9 REGISTRY
CN L-Arginine, L-serylglycyl-L-threonyl-L-valylglycyl- (9CI) (CA INDEX NAME)
OTHER NAMES:
CN PN: WO9954350 SEQID: 1 claimed sequence
SQL 6
SQL 6

SEQ 1 SGTVGR
======
HITS AT: 1-6

REFERENCE 1: 131:332121

```
=> d stat que
L1           62 SEA FILE=REGISTRY ABB=ON  PLU=ON  SGTVGR|SKLMYD/SQSP
L2           59 SEA FILE=REGISTRY ABB=ON  PLU=ON  SGTVGR/SQSP
L3           16 SEA FILE=REGISTRY ABB=ON  PLU=ON  SKLMYD/SQSP
L4           25 SEA FILE=REGISTRY ABB=ON  PLU=ON  L2 AND SQL=<30
L5            3 SEA FILE=HCAPLUS ABB=ON   PLU=ON  L3
L6            3 SEA FILE=HCAPLUS ABB=ON   PLU=ON  L4
L7            3 SEA FILE=HCAPLUS ABB=ON   PLU=ON  L5 OR L6
L9           33 SEA FILE=REGISTRY ABB=ON  PLU=ON  L1 NOT (L3 OR L4)
L10          21 SEA FILE=HCAPLUS ABB=ON   PLU=ON  L9 NOT L7


=>
=>

=> d  ibib abs hitrn l10 1-21

L10  ANSWER 1 OF 21  HCAPLUS  COPYRIGHT 2003 ACS
ACCESSION NUMBER:         2002:781492  HCAPLUS
DOCUMENT NUMBER:          138:1096
TITLE:                    Essential genes in microorganisms and their use as
                          targets for antisense inhibition of proliferation and
                          antibiotic screening
INVENTOR(S):              Wang, Liangus; Zamudio, Carlos; Malone, Cheryl;
                          Haselbeck, Robert; Ohlsen, Kari L.; Zyskind, Judith
                          W.; Wall, Daniel; Trawick, John D.; Carr, Grant J.;
                          Yamamoto, Robert; Forsyth, R. Allyn; Xu, H. Howard
PATENT ASSIGNEE(S):       Elitra Pharmaceuticals, Inc., USA
SOURCE:                   PCT Int. Appl., 1766 pp.
                          CODEN: PIXXD2
DOCUMENT TYPE:            Patent
LANGUAGE:                 English
FAMILY ACC. NUM. COUNT:   22
PATENT INFORMATION:

     PATENT NO.          KIND  DATE          APPLICATION NO.  DATE
     ------------------  ----  --------      ---------------  --------
     WO 2002077183       A2    20021003      WO 2002-XO9107   20020321
         W:  AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN,
             CO, CR, CU, CZ, DE, DE, DK, DK, DM, DZ, EC, EE, EE, ES, FI, FI,
             GB, GD, GE, GH, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ,
             LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO,
             NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SK, SL, TJ, TM,
             TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW, AM, AZ, BY, KG
         RW: GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW, AT, BE, CH,
             CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR,
             BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG
     US 2002061569       A1    20020523      US 2001-815242   20010321
     WO 2002077183       A2    20021003      WO 2002-US9107   20020321
         W:  AE, AG, AL, AM, AT, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH,
             CN, CO, CR, CU, CZ, CZ, DE, DE, DK, DK, DM, DZ, EC, EE, EE, ES,
             FI, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG,
             KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW,
             MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SK,
             SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW,
             AM, AZ, BY, KG
         RW: GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW, AT, BE, CH,
             CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR,
             BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG
PRIORITY APPLN. INFO.:                  US 2001-815242   A  20010321
                                        US 2001-948993   A  20010906
                                        US 2001-342923P  P  20011025
```

US	2002-72851	A	20020208
US	2002-362699P	P	20020306
WO	2002-US9107	A	20020321
US	2000-191078P	P	20000321
US	2000-206848P	P	20000523
US	2000-207727P	P	20000526
US	2000-242578P	P	20001023
US	2000-253625P	P	20001127
US	2000-257931P	P	20001222
US	2001-269308P	P	20010216

AB The sequences of antisense nucleic acids which inhibit the proliferation of prokaryotes are disclosed. Thus, 6213 nucleic acid fragments are identified for which expression inhibits proliferation or is required for proliferation in Enterococcus faecalis, Escherichia coli, Klebsiella pneumoniae, Pseudomonas aeruginosa, Salmonella typhimurium, and Staphylococcus aureus. Cell-based assays which employ the antisense nucleic acids to identify and develop antibiotics are also disclosed. The antisense nucleic acids can also be used to identify proteins required for proliferation, express these proteins or portions thereof, obtain antibodies capable of specifically binding to the expressed proteins, and to use those expressed proteins as a screen to isolate candidate mols. for rational drug discovery programs. The nucleic acids can also be used to screen for homologous nucleic acids that are required for proliferation in cells other than Staphylococcus aureus, Salmonella typhimurium, Klebsiella pneumoniae, and Pseudomonas aeruginosa. The invention provides 38,184 such proliferation-required gene sequences (plus their encoded protein sequences). The nucleic acids of the present invention can also be used in various assay systems to screen for proliferation required genes in other organisms. [This abstr. record is one of twenty records for this document necessitated by the large no. of index entries required to fully index the document and publication system constraints.].

IT **477113-05-6**
RL: BSU (Biological study, unclassified); BUU (Biological use, unclassified); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); USES (Uses)
(amino acid sequence; essential genes in microorganisms and their use as targets for antisense inhibition of proliferation and antibiotic screening)

L10 ANSWER 2 OF 21 HCAPLUS COPYRIGHT 2003 ACS

ACCESSION NUMBER:	2002:723249 HCAPLUS
DOCUMENT NUMBER:	137:227411
TITLE:	Whole-genome comparison of Mycobacterium tuberculosis clinical and laboratory strains
AUTHOR(S):	Fleischmann, R. D.; Alland, D.; Eisen, J. A.; Carpenter, L.; White, O.; Peterson, J.; DeBoy, R.; Dodson, R.; Gwinn, M.; Haft, D.; Hickey, E.; Kolonay, J. F.; Nelson, W. C.; Umayam, L. A.; Ermolaeva, M.; Salzberg, S. L.; Delcher, A.; Utterback, T.; Weidman, J.; Khouri, H.; Gill, J.; Mikula, A.; Bishai, W.; Jacobs, W. R., Jr.; Venter, J. C.; Fraser, C. M.
CORPORATE SOURCE:	The Institute for Genomic Research, Rockville, MD, 20850, USA
SOURCE:	Journal of Bacteriology (2002), 184(19), 5479-5490 CODEN: JOBAAY; ISSN: 0021-9193
PUBLISHER:	American Society for Microbiology
DOCUMENT TYPE:	Journal
LANGUAGE:	English

AB Virulence and immunity are poorly understood in Mycobacterium tuberculosis. The complete genome of the M. tuberculosis clin. strain CDC1551 was sequences and a whole-genome comparison with the lab. strain H37Rv performed in order to identify polymorphic sequences with potential relevance to disease pathogenesis, immunity, and evolution.

Large-sequence and single-nucleotide polymorphisms were found in numerous genes. Polymorphic loci included a phospholipase C, a membrane lipoprotein, members of an adenylate cyclase gene family, and members of the PE/PPE gene family, some of which have been implicated in virulence or the host immune response. Several gene families, including the PE/PPE gene family, also had significantly higher synonymous and nonsynonymous substitution frequencies compared to the genome as a whole. A large sample of M. tuberculosis clin. isolates was tested for a subset of the large-sequence and single-nucleotide polymorphisms and widespread genetic variability was found at many of these loci. Phylogenetic and epidemiol. anal. was carried out to investigate the evolutionary relationships among isolates and the origins of specific polymorphic loci. A no. of these polymorphisms appear to have occurred multiple times as independent events, suggesting that these changes may be under selective pressure. Together, these results demonstrate that polymorphisms among M. tuberculosis strains are more extensive than initially anticipated, and genetic variation may have an important role in disease pathogenesis and immunity. The sequence of the clin. strain CDC1551 of M. tuberculosis was deposited in GenBank/EMBL/DDBJ under accession no. AE000516, and the sequence of the genome of the M. tuberculosis lab. strain H37Rv was recently sequenced and deposited as NC_000962.

IT **457721-02-7**
RL: BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study)
(amino acid sequence; whole-genome comparison of Mycobacterium tuberculosis clin. and lab. strains)

REFERENCE COUNT: 42 THERE ARE 42 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L10 ANSWER 3 OF 21 HCAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 2002:700127 HCAPLUS
DOCUMENT NUMBER: 137:211759
TITLE: Complete genome structure of the thermophilic cyanobacterium Thermosynechococcus elongatus BP-1
AUTHOR(S): Nakamura, Yasukazu; Kaneko, Takakazu; Sato, Shusei; Ikeuchi, Masahiko; Katoh, Hiroshi; Sasamoto, Shigemi; Watanabe, Akiko; Iriguchi, Mayumi; Kawashima, Kumiko; Kimura, Takaharu; Kishida, Yoshie; Kiyokawa, Chiaki; Kohara, Mitsuyo; Matsumoto, Midori; Matsuno, Ai; Nakazaki, Naomi; Shimpo, Sayaka; Sugimoto, Masako; Takeuchi, Chie; Yamada, Manabu; Tabata, Satoshi
CORPORATE SOURCE: Kazusa DNA Research Institute, Chiba, 292-0812, Japan
SOURCE: DNA Research (2002), 9(4), 123-130
CODEN: DARSE8; ISSN: 1340-2838
PUBLISHER: Universal Academy Press
DOCUMENT TYPE: Journal
LANGUAGE: English

AB The entire genome of a thermophilic unicellular cyanobacterium, Thermosynechococcus elongatus BP-1, was sequenced. The genome consisted of a circular chromosome 2,593,857 bp long, and no plasmid was detected. A total of 2475 potential protein-encoding genes, one set of rRNA genes, 42 tRNA genes representing 42 tRNA species and 4 genes for small structural RNAs were assigned to the chromosome by similarity search and computer prediction. The translated products of 56% of the potential protein-encoding genes showed sequence similarity to exptl. identified and predicted proteins of known function, and the products of 34% of these genes showed sequence similarity to the translated products of hypothetical genes. The remaining 10% lacked significant similarity to genes for predicted proteins in the public DNA databases. Sixty-three percent of the T. elongatus genes showed significant sequence similarity to those of both Synechocystis sp. PCC 6803 and Anabaena sp. PCC 7120, while 22% of the genes were unique to this species, indicating a high degree of divergence of the gene information among cyanobacterial strains.

The lack of genes for typical fatty acid desaturases and the presence of more genes for heat-shock proteins in comparison with other mesophilic cyanobacteria may be genomic features of thermophilic strains. A remarkable feature of the genome is the presence of 28 copies of group II introns, 8 of which contained a presumptive gene for maturase/reverse transcriptase. A trace of genome rearrangement mediated by the group II introns was also obsd. The sequence data have been deposited in DDBJ/GenBank/EMBL under accession no. BA000039.

IT **455354-69-5**
RL: BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study)
(amino acid sequence; complete genome structure of the thermophilic cyanobacterium Thermosynechococcus elongatus BP-1)

REFERENCE COUNT: 34 THERE ARE 34 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L10 ANSWER 4 OF 21 HCAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 2002:280980 HCAPLUS
DOCUMENT NUMBER: 137:28849
TITLE: Functional annotation of a full-length Arabidopsis cDNA collection
AUTHOR(S): Seki, Motoaki; Narusaka, Mari; Kamiya, Asako; Ishida, Junko; Satou, Masakazu; Sakurai, Tetsuya; Nakajima, Maiko; Enju, Akiko; Akiyama, Kenji; Oono, Youkò; Muramatsu, Masami; Hayashizaki, Yoshihide; Kawai, Jun; Carninci, Piero; Itoh, Masayoshi; Ishii, Yoshiyuki; Arakawa, Takahiro; Shibata, Kazuhiro; Shinagawa, Akira; Shinozaki, Kazuo
CORPORATE SOURCE: Plant Mutation Exploration Team, Plant Functional Genomics Res. Group, RIKEN Genomic Sciences Center (GSC), 3-1-1 Koyadai, Tsukuba, 305-0074, Japan
SOURCE: Science (Washington, DC, United States) (2002), 296(5565), 141-145
CODEN: SCIEAS; ISSN: 0036-8075
PUBLISHER: American Association for the Advancement of Science
DOCUMENT TYPE: Journal
LANGUAGE: English

AB Full-length cDNAs are essential for the correct annotation of genomic sequences and for the functional anal. of genes and their products. About 155,144 RIKEN Arabidopsis full-length (RAFL) cDNA clones were isolated. The 3'-end expressed sequence tags (ESTs) of 155,144 RAFL cDNAs were clustered into 14,668 nonredundant cDNA groups, about 60% of predicted genes. 5'-ESTs were also obtained from 14,034 nonredundant cDNA groups and a promoter database constructed. The sequence database of the RAFL cDNAs is useful for promoter anal. and correct annotation of predicted transcription units and gene products. Furthermore, the full-length cDNAs are useful resources for analyses of the expression profiles, functions, and structures of plant proteins. [This abstr. record is one of sixteen records for this document necessitated by the large no. of index entries required to fully index the document and publication system constraints.].

IT **437143-91-4**
RL: BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study)
(amino acid sequence; functional annotation of a full-length Arabidopsis cDNA collection)

REFERENCE COUNT: 27 THERE ARE 27 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L10 ANSWER 5 OF 21 HCAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 2002:173242 HCAPLUS
DOCUMENT NUMBER: 136:396934
TITLE: Reagents and kits, such as nucleic acid arrays, for detecting the expression of over 10,000 Drosophila

genes
INVENTOR(S): Venter, J. Craig; Adams, Mark; Li, Peter W. D.; Myers, Eugene W.
PATENT ASSIGNEE(S): PE Corporation (NY), USA
SOURCE: PCT Int. Appl., 21 pp.
CODEN: PIXXD2
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 10
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2001071042	A2	20010927	WO 2001-XI9231	20010323
W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM				
RW: GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG				
WO 2001071042	A2	20010927	WO 2001-US9231	20010323
W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM				
RW: GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG				

PRIORITY APPLN. INFO.:

US 2000-191637P	P	20000323
US 2000-614150	A	20000711
WO 2001-US9231	A	20010323

AB The present invention is based on the sequencing and assembly of the Drosophila melanogaster genome. The present invention provides the primary nucleotide sequence of a large portion of the Drosophila melanogaster genome in a series of genomic and predicted transcript sequences. This information is provided in the form of genomic, transcript and protein sequence information and can be used to generate nucleic acid detection reagents and kits such as nucleic acid arrays. Primary sequences are provided as contiguous strings in a computer-readable format and recorded on media such as floppy disks, hard disks, magnetic tape, CD-ROM, RAM, ROM and hybrids of these categories. Genes/exons can be predicted,sequences can be edited and homol. searches of target motifs can be conducted. [This abstr. record is one of ten records for this document necessitated by the large no. of index entries required to fully index the document and publication system constraints.].

IT **431155-89-4**
RL: ANT (Analyte); BSU (Biological study, unclassified); BUU (Biological use, unclassified); PRP (Properties); ANST (Analytical study); BIOL (Biological study); USES (Uses)
(amino acid sequence; reagents and kits, such as nucleic acid arrays, for detecting the expression of over 10,000 Drosophila genes)

L10 ANSWER 6 OF 21 HCAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 2002:151561 HCAPLUS
DOCUMENT NUMBER: 136:197195
TITLE: Rg2B protein of lettuce in conferring powdery mildew resistance to plants
INVENTOR(S): Michelmore, Richard W.; Shen, Kathy A.; Meyers, Blake

C.
PATENT ASSIGNEE(S): The Regents of the University of California, USA
SOURCE: U.S., 209 pp., Cont.-in-part of U.S. Ser. No. 781,734, abandoned.
CODEN: USXXAM
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 2
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
US 6350933	B1	20020226	US 1998-4838	19980109

PRIORITY APPLN. INFO.: US 1997-781734 B2 19970110
AB RG2B (resistance gene) and RG2B proteins of lettuce which confer powdery mildew disease resistance to plants are provided. The nucleic acids can be used to produce transgenic plants resistant to pests.
IT **401054-29-3**
RL: PRP (Properties)
(unclaimed protein sequence; rg2B protein of lettuce in conferring powdery mildew resistance to plants)
REFERENCE COUNT: 13 THERE ARE 13 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L10 ANSWER 7 OF 21 HCAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 2001:713514 HCAPLUS
DOCUMENT NUMBER: 135:268119
TITLE: Transgenic plants containing heat shock protein Hsp100 and its uses in increasing thermo tolerance of plants and generating products
INVENTOR(S): Lindquist, Susan; Queitsch, Christine; Vierling, Elizabeth
PATENT ASSIGNEE(S): Arch Development Corporation, USA
SOURCE: PCT Int. Appl., 91 pp.
CODEN: PIXXD2
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 1
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2001070929	A2	20010927	WO 2001-US8836	20010320
WO 2001070929	A3	20020314		
W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM				
RW: GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG				
AU 2001047587	A5	20011003	AU 2001-47587	20010320
US 2002053097	A1	20020502	US 2001-812350	20010320

PRIORITY APPLN. INFO.: US 2000-190769P P 20000320
US 2000-198116P P 20000418
WO 2001-US8836 W 20010320
AB A transgenic plant having increased stress tolerance, such as thermo tolerance, comprises a Hsp 100 family nucleic acid sequence. The invention is also directed to methods of producing products from transgenic Hsp 100 plants. Successful use of this method has been demonstrated in cereal, grass, an ornamental plant, a crop plant, a food

plant, an oil-producing plant, a synthetic product-producing plant, an environmental waste-absorbing plant, an alc. producing plant, a medicinal plant, a recreational plant and an animal feed plant.

IT **362648-62-2**
RL: BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study)
(amino acid sequence; transgenic plants contg. heat shock protein Hsp100 and its uses in increasing thermo tolerance of plants and generating products)

L10 ANSWER 8 OF 21 HCAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 2001:164941 HCAPLUS
DOCUMENT NUMBER: 134:188852
TITLE: Massive gene decay in the leprosy bacillus
AUTHOR(S): Cole, S. T.; Eigimeier, K.; Parkhill, J.; James, K. D.; Thomson, N. R.; Wheeler, P. R.; Honore, N.; Garnier, T.; Churcher, C.; Harris, D.; Mungall, K.; Basham, D.; Brown, D.; Chillingworth, T.; Connor, R.; Davies, R. M.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlim, N.; Holroyd, S.; Hornsby, T.; Jagels, K.; Lacroix, C.; Maclean, J.; Moule, S.; Murphy, L.; Oliver, K.; Quail, M. A.; Rajandream, M.-A.; Rutherford, K. M.; Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; Stevens, K.; Taylor, K.; Whitehead, S.; Woodward, J. R.; Barrell, B. G.
CORPORATE SOURCE: Unite de Genetique Moleculaire Bacterienne, Institut Pasteur, Paris, 75724, Fr.
SOURCE: Nature (London, United Kingdom) (2001), 409(6823), 1007-1011
CODEN: NATUAS; ISSN: 0028-0836
PUBLISHER: Nature Publishing Group
DOCUMENT TYPE: Journal
LANGUAGE: English

AB Leprosy, a chronic human neurol. disease, results from infection with the obligate intracellular pathogen Mycobacterium leprae, a close relative fo the tubercle bacillus. Mycobacterium leprae has the longest doubling time of all known bacteria and has thwarted every effort at culture in the lab. Comparing the 3.27-megabase (Mb) genome sequence of an armadillo-derived Indian isolate of the leprosy bacillus with that of Mycobacterium tuberculosis (4.41 Mb) provides clear explanations for these properties and reveals an extreme case of reductive evolution. Less than half of the genome contains functional genes but pseudogenes, with intact counterpart in M tuberculosis, abound. Genome downsizing and the current mosaic arrangement appear to have resulted from extensive recombination events between dispersed repetitive sequences. Gene deletion and decay have eliminated many important metabolic activities including siderophore prodn., part of the oxidative and most of the microaerophilic and anaerobic respiratory chains, and numerous catabolic systems and their regulatory circuits.

IT **169875-44-9**
RL: BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study)
(amino acid sequence; complete genome sequence of Mycobacterium leprae indicates massive gene decay in the leprosy bacillus)

REFERENCE COUNT: 50 THERE ARE 50 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L10 ANSWER 9 OF 21 HCAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 2000:613167 HCAPLUS
DOCUMENT NUMBER: 133:218310
TITLE: DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae

AUTHOR(S): Heidelberg, John F.; Elsen, Jonathan A.; Nelson, William C.; Clayton, Rebecca A.; Gwinn, Michelle L.; Dodson, Robert J.; Haft, Daniel H.; Hickey, Erin K.; Peterson, Jeremy D.; Umayam, Lowell; Gill, Steven R.; Nelson, Karen E.; Read, Timothy D.; Tettelin, Herve; Richardson, Delwood; Ermolaeva, Maria D.; Vamathevan, Jessica; Bass, Steven; Qin, Haiying; Dragoi, Loana; Sellers, Patrick; McDonald, Lisa; Utterback, Teresa; Fleishmann, Robert D.; Nierman, William C.; White, Owen; Salzberg, Steven L.; Smith, Hamilton O.; Colwell, Rita R.; Mekalanos, John J.; Venter, J. Craig; Fraser, Claire M.

CORPORATE SOURCE: The Institute for Genomic Research, Rockville, MD, 20850, USA

SOURCE: Nature (London) (2000), 406(6795), 477-483
CODEN: NATUAS; ISSN: 0028-0836

PUBLISHER: Nature Publishing Group

DOCUMENT TYPE: Journal

LANGUAGE: English

AB The complete genomic sequence of the gram-neg., .gamma.-Proteobacterium Vibrio cholerae El Tor N16961 was detd. to be 4,033,460 bp. The genome consists of two circular chromosomes of 2,961,146 bp and 1,072,314 bp that together encode 3885 open reading frames. The vast majority of recognizable genes for essential cell functions (such as DNA replication, transcription, translation, and cell-wall biosynthesis) and pathogenicity (for example, toxins, surface antigens, and adhesins) are located on the large chromosome. In contrast, the small chromosome contains a larger fraction (59%) of hypothetical genes compared with the large chromosome (42%), and also contains many more genes that appear to have origins other than the .gamma.-Proteobacteria. The small chromosome also carries a gene capture system (the integron island) and host 'addiction' genes that are typically found on plasmids; thus, the small chromosome may have originally been a megaplasmid that was captured by an ancestral Vibrio species. The V. cholerae genomic sequence provides a starting point for understanding how a free-living, environmental organism emerged to become a significant human bacterial pathogen.

IT **290395-99-2**
RL: BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study)
(amino acid sequence; DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae)

REFERENCE COUNT: 50 THERE ARE 50 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L10 ANSWER 10 OF 21 HCAPLUS COPYRIGHT 2003 ACS

ACCESSION NUMBER: 2000:246848 HCAPLUS

DOCUMENT NUMBER: 132:289494

TITLE: The genome sequence of Drosophila melanogaster

AUTHOR(S): Adams, Mark D.; Celniker, Susan E.; Holt, Robert A.; Evans, Cheryl A.; Gocayne, Jeannine D.; Amanatides, Peter G.; Scherer, Steven E.; Li, Peter W.; Hoskins, Roger A.; Galle, Richard F.; George, Reed A.; Lewis, Suzanna E.; Richards, Stephen; Ashburner, Michael; Henderson, Scott N.; Sutton, Granger G.; Wortman, Jennifer R.; Yandell, Mark D.; Zhang, Qing; Chen, Lin X.; Brandon, Rhonda C.; Rogers, Yu-Hui C.; Blazej, Robert G.; Champe, Mark; Pfeiffer, Barret D.; Wan, Kenneth H.; Doyle, Clare; Baxter, Evan G.; Helt, Gregg; Nelson, Catherine R.; Miklos, George L. Gabor; Abril, Josep F.; Agbayani, Anna; An, Hui-Jin; Andrews-Pfannkoch, Cynthia; Baldwin, Danita; Ballew, Richard M.; Basu, Anand; Baxendale, James; Bayraktaroglu, Leyla; Beasley, Ellen M.; Beeson, Karen

Y.; Benos, P. V.; Berman, Benjamin P.; Bhandari, Deepali; Bolshakov, Slava; Borkova, Dana; Botchan, Michael R.; Bouck, John; Brokstein, Peter; Brottier, Phillipe; Burtis, Kenneth C.; Busam, Dana A.; Butler, Heather; Cadieu, Edouard; Center, Angela; Chandra, Ishwar; Cherry, J. Michael; Cawley, Simon; Dahlke, Carl; Davenport, Lionel B.; Davies, Peter; De Pablos, Beatriz; Delcher, Arthur; Deng, Zuoming; Mays, Anne Deslattes; Dew, Ian; Dietz, Suzanne M.; Dodson, Kristina; Doup, Lisa E.; Downes, Michael; Dugan-Rocha, Shannon; Dunkov, Boris C.; Dunn, Patrick; Durbin, Kenneth J.; Evangelista, Carlos C.; Ferraz, Concepcion; Ferriera, Steven; Fleischmann, Wolfgang; Foster, Carl; Gabrielian, Andrei E.; Garg, Neha S.; Gelbart, William M.; Glasser, Ken; Glodek, Anna; Gong, Fangcheng; Gorrell, J. Harley; Gu, Zhiping; Guan, Ping; Harris, Michael; Harris, Nomi L.; Harvey, Damon; Heiman, Thomas J.; Hernandez, Judith R.; Houck, Jarrett; Hostin, Damon; Houston, Kathryn A.; Howland, Timothy J.; Wei, Ming-Hui; Ibegwam, Chinyere; Jalali, Mena; Kalush, Francis; Karpen, Gary H.; Ke, Zhaoxi; Kennison, James A.; Ketchum, Karen A.; Kimmel, Bruce E.; Kodira, Chinnappa D.; Kraft, Cheryl; Kravitz, Saul; Kulp, David; Lai, Zhongwu; Lasko, Paul; Lei, Yiding; Levitsky, Alexander A.; Li, Jiayin; Li, Zhenya; Liang, Yong; Lin, Xiaoying; Liu, Xiangjun; Mattei, Bettina; McIntosh, Tina C.; McLeod, Michael P.; McPherson, Duncan; Merkulov, Gennady; Milshina, Natalia V.; Mobarry, Clark; Morris, Joe; Moshrefi, Ali; Mount, Stephen M.; Moy, Mee; Murphy, Brian; Murphy, Lee; Muzny, Donna M.; Nelson, David L.; Nelson, David R.; Nelson, Keith A.; Nixon, Katherine; Nusskern, Deborah R.; Pacleb, Joanne M.; Palazzolo, Michael; Pittman, Gjange S.; Pan, Sue; Pollard, John; Puri, Vinita; Reese, Martin G.; Reinert, Knut; Remington, Karin; Saunders, Robert D. C.; Scheeler, Frederick; Shen, Hua; Shue, Bixiang Christopher; Siden-Kiamos, Inga; Simpson, Michael; Skupski, Marian P.; Smith, Tom; Spier, Eugene; Spradling, Allan C.; Stapleton, Mark; Strong, Renee; Sun, Eric; Svirskas, Robert; Tector, Cyndee; Turner, Russell; Venter, Eli; Wang, Aihui H.; Wang, Xin; Wang, Zhen-Yuan; Wassarman, David A.; Weinstock, George M.; Weissenbach, Jean; Williams, Sherita M.; Woodage, Trevor; Worley, Kim C.; Wu, David; Yang, Song; Yao, Q. Alison; Ye, Jane; Yeh, Ru-Fang; Zaveri, Jayshree S.; Zhan, Ming; Zhang, Guangren; Zhao, Qi; Zheng, Liansheng; Zheng, Xiangqun H.; Zhong, Fei N.; Zhong, Wenyan; Zhou, Xiaojun; Zhu, Shiaoping; Zhu, Xiaohong; Smith, Hamilton O.; Gibbs, Richard A.; Myers, Eugene W.; Rubin, Gerald M.; Venter, J. Craig

CORPORATE SOURCE: Celera Genomics, Rockville, MD, 20850, USA

SOURCE: Science (Washington, D. C.) (2000), 287(5461), 2185-2195
CODEN: SCIEAS; ISSN: 0036-8075

PUBLISHER: American Association for the Advancement of Science
DOCUMENT TYPE: Journal
LANGUAGE: English

AB The fly Drosophila melanogaster is one of the most intensively studied organisms in biol. and serves as a model system for the investigation of many developmental and cellular processes common to higher eukaryotes, including humans. The nucleotide sequence was detd. of nearly all of the .apprx.120-megabase euchromatic portion of the Drosophila genome using a

whole-genome shotgun sequencing strategy supported by extensive clone-based sequence and a high-quality bacterial artificial chromosome phys. map. Efforts are under way to close the remaining gaps; however, the sequence is of sufficient accuracy and contiguity to be declared substantially complete and to support an initial anal. of genome structure and preliminary gene annotation and interpretation. The genome encodes .apprx.13,600 genes, somewhat fewer than the smaller Caenorhabditis elegans genome, but with comparable functional diversity. Access to supporting information on each gene is available through FlyBast at http://flybase.bio.indiana.edu and through Celera at www.celera.com; the sequences are deposited in GenBank with Accession Nos. AE002566-AE003403. [This abstr. record is one of 4 records for this document necessitated by the large no. of index entries required to fully index the document and publication system constraints.].

IT **263502-19-8**
RL: BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study)
(amino acid sequence; genome sequence of Drosophila melanogaster)

L10 ANSWER 11 OF 21 HCAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 1998:773005 HCAPLUS
DOCUMENT NUMBER: 130:120325
TITLE: Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. [Erratum to document cited in CA129:77224]
AUTHOR(S): Cole, S. T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S. V.; Eiglmeier, K.; Gas, S.; Barry, C. E., III; Tekaia, F.; Badcock, K.; Basham, D.; Brown, D.; Chillingworth, T.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Hornsby, T.; Jagels, K.; Krogh, A.; McLean, J.; Moule, S.; Murphy, L.; Oliver, K.; Osborne, J.; Quail, M. A.; Rajandream, M.-A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, J.; Squares, R.; Squares, S.; Sulston, J. E.; Taylor, K.; Whitehead, S.; Barrell, B. G.
CORPORATE SOURCE: Sanger Cent., Wellcome Trust Genome Campus, Hinxton, CB10 1SA, UK
SOURCE: Nature (London) (1998), 396(6707), 190-198
CODEN: NATUAS; ISSN: 0028-0836
PUBLISHER: Macmillan Magazines
DOCUMENT TYPE: Journal
LANGUAGE: English

AB Table 1 was published with some symbols missing; the correct version can be found at http://www.sanger.ac.uk and is given here. In Fig. 2, Rv0649 was incorrectly labeled as fadD37 instead of fabD2. Two of the genes for mycolyl transferases were inverted: Rv0129c encodes antigen 85C and not 85c' as stated, whereas Rv3803c codes for the secreted protein MPT51 and not antigen 85C (Infect. Immun. 59, 372-382; 1991); Rv3803c is now designated fbpD. The sequence of Rv0746 from M. bovis BCG-Pasteur presented in Fig. 5 b was incorrect and should have shown a 16-codon deletion instead of 29.

IT **208861-47-6**
RL: PRP (Properties)
(deciphering the biol. of Mycobacterium tuberculosis from the complete genome sequence (Erratum))

L10 ANSWER 12 OF 21 HCAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 1998:490477 HCAPLUS
DOCUMENT NUMBER: 129:145635
TITLE: Isolation, sequence, cloning, and expression of lettuce Rg genes and proteins for conferring disease resistance to plants

INVENTOR(S): Shen, Kathy; Meyers, Blake; Michelmore, Richard W.
PATENT ASSIGNEE(S): The Regents of the University of California, USA
SOURCE: PCT Int. Appl., 182 pp.
CODEN: PIXXD2
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 2
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 9830083	A1	19980716	WO 1998-US615	19980109
W: CA, JP				
RW: AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE				
EP 969714	A1	20000112	EP 1998-902515	19980109
R: FR, GB, NL, SE				

PRIORITY APPLN. INFO.: US 1997-781734 A 19970110
WO 1998-US615 W 19980109

AB The present invention provides RG (resistance gene) genes fro Lactuca sativa that belong to RG1, RG2, RG3, RG4, RG5, RG6, and RG7 families, and which encode proteins which confer disease resistance to plants, particularly transgenic lettuce. The RG proteins contain a leucine-rich region and/or nucleotide-binding site. The nucleic acids can be used to produce transgenic various plants resistant to pests. Preferred plants are Lactuca sativa, and subspecies crispa, longifolia, and asparagina. The RG genes can be used in marker-aided selection, and partial sequences used as hybridization probes or as the basis for oligonucleotide primers to amplify nucleic acids by PCR. Transgenic plants can also be constructed using RG promoters. Antibodies to proteins of the invention are also provided. Antisense oligonucleotides are also constructed which are capable of binding RG message which can inhibit RG activity by targeting mRNA.

IT **210885-91-9**, Protein (Lactuca sativa gene RG1A)
RL: AGR (Agricultural use); BAC (Biological activity or effector, except adverse); BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study); USES (Uses)
(amino acid sequence; isolation, sequence, cloning, and expression of lettuce Rg genes and proteins for conferring disease resistance to plants)

REFERENCE COUNT: 5 THERE ARE 5 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L10 ANSWER 13 OF 21 HCAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 1998:389708 HCAPLUS
DOCUMENT NUMBER: 129:77224
TITLE: Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence
AUTHOR(S): Cole, S. T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S. V.; Eiglmeier, K.; Gas, S.; Barry, C. E., III.; Tekaia, F.; Badcock, K.; Basham, D.; Brown, D.; Chillingworth, T.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Hornsby, T.; Jagels, K.; Krogh, A.; McLean, J.; Moule, S.; Murphy, L.; Oliver, K.; Osborne, J.; Quail, M. A.; Rajandream, M.-A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, J.; Squares, R.; Squares, S.; Sulston, J. E.; Taylor, K.; Whitehead, S.; Barrell, B. G.
CORPORATE SOURCE: Sanger Cent., Wellcome Trust Genome Campus, Hinxton, CB10 1SA, UK
SOURCE: Nature (London) (1998), 393(6685), 537-544
CODEN: NATUAS; ISSN: 0028-0836
PUBLISHER: Macmillan Magazines

DOCUMENT TYPE: Journal
LANGUAGE: English
AB Countless millions of people have died from tuberculosis, a chronic infectious disease caused by the tubercle bacillus. The complete genome sequence of the best-characterized strain of Mycobacterium tuberculosis, H37Rv, was detd. and analyzed in order to improve our understanding of the biol. of this slow-growing pathogen and to help the conception of new prophylactic and therapeutic interventions. The genome comprises 4,411,529 base pairs, contains around 4000 genes, and has a very high G+C content that is reflected in the biased amino acid content of the proteins. M. tuberculosis differs radically from other bacteria in that a very large portion of its coding capacity is devoted to the prodn. of enzymes involved in lipogenesis and lipolysis, and to 2 new families of glycine-rich proteins with a repetitive structure that may represent a source of antigenic variation.
IT **208861-47-6**
RL: PRP (Properties)
(amino acid sequence; deciphering the biol. of Mycobacterium tuberculosis from the complete genome sequence)
REFERENCE COUNT: 46 THERE ARE 46 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L10 ANSWER 14 OF 21 HCAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 1997:736785 HCAPLUS
DOCUMENT NUMBER: 128:19264
TITLE: Complete genome sequence of Methanobacterium thermoautotrophicum .DELTA.H: functional anal. and comparative genomics
AUTHOR(S): Smith, Douglas R.; Doucette-Stamm, Lynn A.; Deloughery, Craig; Lee, Hongmei; Dubois, Joann; Aldredge, Tyler; Bashirzadeh, Romina; Blakely, Derron; Cook, Robin; Gilbert, Katie; Harrison, Dawn; Hoang, Lieu; Keagle, Pamela; Lumm, Wendy; Pothier, Bryan; Qiu, Dayong; Spadafora, Rob; Vicaire, Rita; Wang, Ying; Wierzbowski, Jamey; Gibson, Rene; Jiwani, Nilofer; Caruso, Anthony; Bush, David; Safer, Hershel; Patwell, Donivan; Prabhakar, Shashi; McDougall, Steve; Shimer, George; Goyal, Anil; Pietrokovski, Shmuel; Church, George M.; Daniels, Charles J.; Mao, Jen-I.; Rice, Phil; Nolling, Jork; Reeve, John N.
CORPORATE SOURCE: Collaborative Research Division, Genome Therapeutics Corporation, Waltham, MA, 02154, USA
SOURCE: Journal of Bacteriology (1997), 179(22), 7135-7155
CODEN: JOBAAY; ISSN: 0021-9193
PUBLISHER: American Society for Microbiology
DOCUMENT TYPE: Journal
LANGUAGE: English
AB The complete 1,751,377-bp sequence of the genome of the thermophilic archaeon Methanobacterium thermoautotrophicum AH was been detd. by a whole-genome shotgun sequencing approach. A total of 1855 open reading frames (ORFs) were identified that appear to encode polypeptides, 844 (46%) of which were assigned putative functions based on their similarities to database sequences with assigned functions. A total of 514 (28%) of the ORF-encoded polypeptides are related to sequences with unknown functions, and 496 (27%) have little or no homol. to sequences in public databases. Comparisons with Eucarya-, Bacteria-, and Archaea-specific databases reveal that 1013 of the putative gene products (54%) are most similar to polypeptide sequences described previously for other organisms in the domain Archaea. Comparisons with the Methanococcus jannaschii genome data underline the extensive divergence that has occurred between these 2 methanogens; only 352 (19%) of M. thermoautotrophicum ORFs encode sequences that are >50% identical to M. jannaschii polypeptides, and there is little conservation in the relative

locations of orthologous genes. When the M. thermoautotrophicum ORFs are compared to sequences from only the eukaryal and bacterial domains, 786 (42%) are more similar to bacterial sequences and 241 (13%) are more similar to eukaryal sequences. The bacterial domain-like gene products include the majority of those predicted to be involved in cofactor and small mol. biosyntheses, intermediary metab., transport, nitrogen fixation, regulatory functions, and interactions with the environment. Most proteins predicted to be involved in DNA metab., transcription, and translation are more similar to eukaryal sequences. Gene structure and organization have features that are typical of the Bacteria, including genes that encode polypeptides closely related to eukaryal proteins. There are 24 polypeptides that could form 2-component sensor kinase-response regulator systems and homologs of the bacterial Hsp70-response proteins DnaK and DnaJ, which are notably absent in M. jannaschii. DNA replication initiation and chromosome packaging in M. thermoautotrophicum are predicted to have eukaryal features, based on the presence of two Cdc6 homologs and three histones; however, the presence of an ftsZ gene indicates a bacterial type of cell division initiation. The DNA polymerases include an X-family repair type and an unusual archaeal B type formed by 2 sep. polypeptides. The DNA-dependent RNA polymerase (RNAP) subunits A', A", B', B" and H are encoded in a typical archaeal RNAP operon, although a second A' subunit-encoding gene is present at a remote location. There are two rRNA operons, and 39 tRNA genes are dispersed around the genome, although most of these occur in clusters. Three of the tRNA genes have introns, including the tRNAPro (GGG) gene, which contains a second intron at an unprecedented location. There is no selenocysteinyl-tRNA gene nor evidence for classically organized IS elements, prophages, or plasmids. The genome contains one intein and 2 extended repeats (3.6 and 8.6 kb) that are members of a family with 18 representatives in the M. jannaschii genome.

IT **198576-75-9**
RL: PRP (Properties)
(amino acid sequence; complete genome sequence of Methanobacterium thermoautotrophicum .DELTA.H with functional anal. and comparative genomics)

L10 ANSWER 15 OF 21 HCAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 1997:556747 HCAPLUS
DOCUMENT NUMBER: 127:273529
TITLE: Multiplex sequencing of 1.5 Mb of the Mycobacterium leprae genome
AUTHOR(S): Smith, Douglas R.; Richterich, Pewter; Rubenfield, Marc; Rice, Philip W.; Butler, Carol; Lee, Hong-Mei; Kirst, Susan; Gundersen, Kristin; Abendschan, Kari; Xu, Qinxue; Chung, Maria; Deloughery, Craig; Aldredge, Tyler; Maher, James; Lundstrom, Ronald; Tulig, Craig; Falls, Kathleen; Imrich, Joan; Torrey, Dana; Engelstein, Marcy; Breton, Gary; Madan, Deepika; Nietupski, Raymond; Seitz, Bruce; Connelly, Steven; McDougall, Steven; Safer, Hershel; Gibson, Rene; Doucette-Stamm, Lynn; Eiglmeier, Karin; Bergh, Staffan; Cole, Stewart T.; Robinson, Keith; Richterich, Laura; Johnson, Jason; Church, George M.; Mao, Jen-i
CORPORATE SOURCE: Collaborative Res. Div., Genome Therapeutics Corp., Waltham, MA, 02154, USA
SOURCE: Genome Research (1997), 7(8), 802-819
CODEN: GEREFS; ISSN: 1088-9051
PUBLISHER: Cold Spring Harbor Laboratory Press
DOCUMENT TYPE: Journal
LANGUAGE: English

AB The nucleotide sequence of 1.5 Mb of genomic DNA from Mycobacterium leprae was detd. using computer-assisted multiplex sequencing technol. This

brings the 2.8-Mb M. leprae genome sequence to ~66% completion. The sequences, derived from 43 recombinant cosmids, contain 1046 putative protein-coding genes, 44 repetitive regions, 3 rRNAs, and 15 tRNAs. The gene d. of one per 1.4 kb is slightly lower than that of Mycoplasma (1.2 kb). Of the protein coding genes, 44% have significant matches to genes with well-defined functions. Comparison of 1157 M. leprae and 1564 Mycobacterium tuberculosis proteins shows a complex mosaic of homologous genomic blocks with up to 22 adjacent proteins in conserved map order. Matches to known enzymic, antigenic, membrane, cell wall, cell division, multidrug resistance, and virulence proteins suggest therapeutic and vaccine targets. Unusual features of the M. leprae genome include large polyketide synthase (pks) operons, inteins, and highly fragmented pseudogenes.

IT **196319-84-3**
RL: BOC (Biological occurrence); BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study); OCCU (Occurrence)
(amino acid sequence; multiplex sequencing of 1.5 Mb of Mycobacterium leprae genome)

L10 ANSWER 16 OF 21 HCAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 1995:715583 HCAPLUS
DOCUMENT NUMBER: 123:307660
TITLE: The Mycobacterium leprae genome: systematic sequence analysis identifies key catabolic enzymes, ATP-dependent transport systems and a novel polA locus associated with genomic variability
AUTHOR(S): Esihi, Hafida; Cole, Stewart T.
CORPORATE SOURCE: Unite Genetique Moleculaire Bacterienne, Inst. Pasteur, Paris, 75724, Fr.
SOURCE: Molecular Microbiology (1995), 16(5), 909-19
CODEN: MOMIEE; ISSN: 0950-382X
PUBLISHER: Blackwell
DOCUMENT TYPE: Journal
LANGUAGE: English
AB In the framework of the mycobacterial genome sequencing project, a continuous 37049 bp sequence from the Mycobacterium leprae chromosome has been detd. Computer anal. revealed 10 complete open reading frames, and nine of their products show similarity to known proteins. Seven of these were identified as the enzyme isocitrate lyase, two P-type ATPase cation transporters, two AMP-binding proteins, the ribosomal protein S1, and DNA polymerase I. Interestingly, the polA gene, encoding DNA polymerase, is flanked by two inverted copies of a new class of the M. leprae specific repetitive sequence, RLEP, and this structure resembles a transposable element. A second copy of this element was found at another locus in the genome, but the two copies were not present in equal amts. and could not be found in all isolates of M. leprae. This is the first evidence for genomic variability in the leprosy bacillus and might ultimately be useful for developing a mol. test capable of distinguishing between strains of M. leprae.

IT **169875-44-9**
RL: PRP (Properties)
(amino acid sequence; sequence anal. of Mycobacterium leprae genome identifies catabolic enzymes, ATP-dependent transport systems and a novel polA locus assocd. with genomic variability)

L10 ANSWER 17 OF 21 HCAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 1995:287964 HCAPLUS
DOCUMENT NUMBER: 123:26572
TITLE: Primary structure of the Streptomyces R61 extracellular DD-peptidase. 1. Cloning into Streptomyces lividans and nucleotide sequence of the gene. [Erratum to document cited in CA106:169914]
AUTHOR(S): Duez, Colette; Piron-Fraipont, Claudine; Joris,

Bernard; Dusart, Jean; Urdea, Mickey S.; Martial, Joseph A.; Frere, Jean Marie; Ghuysen, Jean Marie
CORPORATE SOURCE: Fac. Med., Univ. Liege, Liege, Belg.
SOURCE: European Journal of Biochemistry (1994), 224(3), 1079
CODEN: EJBCAI; ISSN: 0014-2956
PUBLISHER: Springer
DOCUMENT TYPE: Journal
LANGUAGE: English
AB The errors were not reflected in the abstr. but were reflected in the structures of the indexed sequence entries.
IT **107761-90-0**
RL: BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study)
(amino acid sequence of (Erratum))

L10 ANSWER 18 OF 21 HCAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 1994:647466 HCAPLUS
DOCUMENT NUMBER: 121:247466
TITLE: the SPT10 and SPT21 genes of Saccharomyces cerevisiae
AUTHOR(S): Natsoulis, Georges; Winston, Fred; Boeke, Jef D.
CORPORATE SOURCE: Dep. Mol. Biol. Genet., Johns Hopkins Sch. Med., Baltimore, MD, 21205, USA
SOURCE: Genetics (1994), 136(1), 93-105
CODEN: GENTAE; ISSN: 0016-6731
DOCUMENT TYPE: Journal
LANGUAGE: English
AB Mutations in the SPT10 and SPT21 genes were originally isolated as suppressors of Ty and LTR (.delta.) insertion mutations in Saccharomyces cerevisiae, and the genes were shown to be required for normal transcription at a no. of loci in yeast. Now the authors have cloned, sequenced, mapped and mutagenized SPT10 and SPT21. Since the spt10 mutation used to clone SPT10 resulted in very poor transformation efficiency, a novel method making use of the kar 1-1 mutation was used. Neither SPT gene is essential for growth, and constructed null alleles cause phenotypes similar to those caused by spontaneous mutations in the genes. Spt10 null alleles are strong suppressor mutations and cause extremely slow growth. Certain spt10 spontaneous alleles are good suppressors but have a normal growth rate, suggesting that the SPT10 protein may have two distinct functions. An amino acid sequence motif that is similar to the Zn-finger motif was found in SPT10. Mutation of the second Cys residue in this motif resulted in loss of complementation of the suppression phenotype but a normal growth rate. Thus, this motif may reside in a part of the SPT10 protein that is important for transcriptional regulation but not for normal growth. Both the SPT10 and SPT21 proteins are relatively tolerant of large deletions; in both cases deletions of the C-terminus resulted in at least partially functional proteins; also, a large internal deletion in SPT21 was phenotypically wild type.
IT **158650-69-2**
RL: PRP (Properties)
(amino acid sequence)

L10 ANSWER 19 OF 21 HCAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 1994:126392 HCAPLUS
DOCUMENT NUMBER: 120:126392
TITLE: Characterization of cDNA for a dehydration-inducible gene that encodes a Clp A, B-like protein in Arabidopsis thaliana L
AUTHOR(S): Kiyosue, Tomohiro; Yamaguchi-Shinozaki, Kazuko; Shinozaki, Kazuo
CORPORATE SOURCE: Lab. Plant. Mol. Biol., Inst. Phys. Chem. Res., Tsukuba, 305, Japan
SOURCE: Biochemical and Biophysical Research Communications

(1993), 196(3), 1214-20
CODEN: BBRCA9; ISSN: 0006-291X
DOCUMENT TYPE: Journal
LANGUAGE: English
AB The sequence was detd. for a cDNA clone, designated ERD1, isolated from a cDNA library of 1-h-dehydrated plants of Arabidopsis thaliana. The clone (3150 bp) contains an open reading frame of 946 amino acid residues with >34% sequence identity to the regulatory subunit of the Clp ATP-dependent protease in Escherichia coli and contains a putative chloroplast-targeting signal at the N-terminus. Southern blot anal. suggested the presence of addnl. ERD1-related genes in A. thaliana. The expression of ERD1 gene was strongly induced by dehydration-stress but not by heat-, cold-, or heavy-metal-stress. N addn., ERD1 gene expression was not strongly affected by treatment with plant growth regulators, such as auxin, cytokinin, abscisic acid, and gibberellic acid, or by starvation-stress for 10 h.
IT **152889-46-8**
RL: PRP (Properties)
(amino acid sequence and homol. to Clp ATP-dependent protease regulatory subunit of)

L10 ANSWER 20 OF 21 HCAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 1994:1889 HCAPLUS
DOCUMENT NUMBER: 120:1889
TITLE: Defective Marek's disease virus DNA contains a gene encoding a potential nuclear DNA binding protein and a HSV a-like sequence
AUTHOR(S): Camp, Heidi S.; Silva, Robert F.; Coussens, Paul M.
CORPORATE SOURCE: Dep. Anim. Sci., Michigan State Univ., East Lansing, MI, 48824, USA
SOURCE: Virology (1993), 196(2), 484-95
CODEN: VIRLAX; ISSN: 0042-6822
DOCUMENT TYPE: Journal
LANGUAGE: English
AB Four RNA transcripts from chicken embryo fibroblast cells infected with Marek's disease virus (MDV) strain 281MI/1 hybridized to the 4-kbp MDV replicon DNA. In an attempt to identify open reading frame coding for the four transcripts, the authors detd. the nucleotide sequences of 4-kbp replicon DNA (represents a single monomeric repeat unit of defective MDV genome). Computer anal. indicates that the 4-kbp MDV replicon DNA contains two intact open reading frame (ORFs) with common promoter regulatory elements. ORF-A codes for a putative 204 amino acid protein that shares 21 and 36% amino acid sequence identity to nuclear DNA binding proteins such as the EBNA-1 of Epstein-Barr virus and galline, a chicken sperm histone protein, resp. ORF-B encodes for a potential 350 amino acid protein, which did not show any significant amino acid sequence identity to known protein sequences within Swiss-Protein data base. ORF-B may, therefore, encode a MDV specific protein. The 5'-region of MDV replicon DNA revealed seven reiterated copies of an 11-bp motif sharing 8 out of 11 nucleotide sequence identity to DR2 element of the herpes simplex virus strain USA-8 a sequence.
IT **151690-13-0**
RL: PRP (Properties)
(amino acid sequence of)

L10 ANSWER 21 OF 21 HCAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 1987:169914 HCAPLUS
DOCUMENT NUMBER: 106:169914
TITLE: Primary structure of the Streptomyces R61 extracellular DD-peptidase. 1. Cloning into Streptomyces lividans and nucleotide sequence of the gene
AUTHOR(S): Duez, Colette; Piron-Fraipont, Claudine; Joris,

Bernard; Dusart, Jean; Urdea, Mickey S.; Martial, Joseph A.; Frere, Jean Marie; Ghuysen, Jean Marie
CORPORATE SOURCE: Fac. Med., Univ. Liege, Liege, Belg.
SOURCE: European Journal of Biochemistry (1987), 162(3), 509-18
CODEN: EJBCAI; ISSN: 0014-2956
DOCUMENT TYPE: Journal
LANGUAGE: English
AB An 11,450-base DNA fragment contg. the gene for the extracellular active-site serine DD-peptidase [9046-67-7] of Streptomyces R61 was cloned in S. lividans using the high-copy-no. plasmid pIJ702 as vector. Amplified expression of the excreted enzyme was obsd. Producing clones were identified with the help of a specific antiserum directed against the pure DD-peptidase. The coding sequence of the gene was then located by hybridization with a specific nucleotide probe and sub-fragments were obtained from which the nucleotide sequence of the structural gene and the putative promoter and terminator regions were detd. The sequence suggests that the gene codes for a 406-amino-acid precursor. When compared with the excreted mature DD-peptidase, this precursor possesses a cleavable 31-amino-acid N-terminal extension which has the characteristics of a signal peptide, and a cleavable 26-amino-acid C-terminal extension. On the basis of the data of B. Joris et al. (1987), the open reading frame coding for the synthesis of the DD-peptidase was established. Comparison of the primary structure of the Streptomyces R61 DD-peptidase with those of several active-site serine .beta.-lactamases [9073-60-3] and penicillin-binding proteins of Escherichia coli shows homol. in those sequences that comprise the active-site serine residue. When the comparison is broadened to the complete amino acid sequences, significant homol. is obsd. only for the pair Streptomyces R61 DD-peptidase/E. coli ampC .beta.-lactamase (class C). Since the Streptomyces R61 DD-peptidase and .beta.-lactamases of class A have very similar 3-dimensional structures, it is concluded that these tertiary features are probably also shared by the .beta.-lactamases of class C, i.e. that the Streptomyces R61 DD-peptidase and the .beta.-lactamases of classes A and C are related in an evolutionary sense.
IT **107761-90-0**
RL: PRP (Properties)
(amino acid sequence of)

=>
=>

=> fil reg
FILE 'REGISTRY' ENTERED AT 08:55:23 ON 26 FEB 2003

Property values tagged with IC are from the ZIC/VINITI data file provided by InfoChem.

STRUCTURE FILE UPDATES: 24 FEB 2003 HIGHEST RN 494745-03-8
DICTIONARY FILE UPDATES: 24 FEB 2003 HIGHEST RN 494745-03-8

TSCA INFORMATION NOW CURRENT THROUGH MAY 20, 2002

Please note that search-term pricing does apply when conducting SmartSELECT searches.

Crossover limits have been increased. See HELP CROSSOVER for details.

Experimental and calculated property data are now available. See HELP

```
PROPERTIES for more information.  See STNote 27, Searching Properties
in the CAS Registry File, for complete details:
http://www.cas.org/ONLINE/STN/STNOTES/stnotes27.pdf

=>
=>

=> d .seq l9 1-33

L9   ANSWER 1 OF 33  REGISTRY  COPYRIGHT 2003 ACS
RN   489172-01-2  REGISTRY
CN   GenBank AAD48882 (9CI)  (CA INDEX NAME)
OTHER NAMES:
CN   GenBank AAD48882 (Translated from: GenBank U52150)
SQL  358

SEQ    151 AGGVMTAIAV YLLSVKSGTV GRYRLVLTGI GVGSVLTALN GLLLVKGSLD
                           ==== ==
HITS AT:   167-172

**RELATED SEQUENCES AVAILABLE WITH SEQLINK**

L9   ANSWER 2 OF 33  REGISTRY  COPYRIGHT 2003 ACS
RN   488584-20-9  REGISTRY
CN   GenBank AAA62239 (9CI)  (CA INDEX NAME)
OTHER NAMES:
CN   GenBank AAA62239 (Translated from: GenBank M26842)
SQL  406

SEQ      1 MVSGTVGRGT ALGAVLLALL AVPAQAGTAA AADLPAPDDT GLQAVLHTAL
             ======
HITS AT:   3-8

L9   ANSWER 3 OF 33  REGISTRY  COPYRIGHT 2003 ACS
RN   488219-71-2  REGISTRY
CN   GenBank CAA28756 (9CI)  (CA INDEX NAME)
OTHER NAMES:
CN   GenBank CAA28756 (Translated from: GenBank X05109)
SQL  406

SEQ      1 MVSGTVGRGT ALGAVLLALL AVPAQAGTAA AADLPAPDDT GLQAVLHTAL
             ======
HITS AT:   3-8

**RELATED SEQUENCES AVAILABLE WITH SEQLINK**

L9   ANSWER 4 OF 33  REGISTRY  COPYRIGHT 2003 ACS
RN   487546-43-0  REGISTRY
CN   GenBank CAB97254 (9CI)  (CA INDEX NAME)
OTHER NAMES:
CN   GenBank CAB97254 (Translated from: GenBank X55810)
SQL  91

SEQ      1 MVSGTVGRGT ALGAVLLALL AVPAQAGTAA AADLPAPDDT GLQAVLHTAL
             ======
HITS AT:   3-8

L9   ANSWER 5 OF 33  REGISTRY  COPYRIGHT 2003 ACS
RN   487225-76-3  REGISTRY
CN   GenBank AAA35078 (9CI)  (CA INDEX NAME)
OTHER NAMES:
CN   GenBank AAA35078 (Translated from: GenBank L24436)
SQL  758
```

SEQ 251 NNTNSGTVGR RQTNPMPAPK AVRTQSLPIW NLKPNIANTG FPRNSIAHKI
======
HITS AT: 255-260

RELATED SEQUENCES AVAILABLE WITH SEQLINK

L9 ANSWER 6 OF 33 REGISTRY COPYRIGHT 2003 ACS
RN 486828-06-2 REGISTRY
CN GenBank CAA89912 (9CI) (CA INDEX NAME)
OTHER NAMES:
CN GenBank CAA89912 (Translated from: GenBank Z49808)
SQL 758

SEQ 251 NNTNSGTVGR RQTNPMPAPK AVRTQSLPIW NLKPNIANTG FPRNSIAHKI
======
HITS AT: 255-260

RELATED SEQUENCES AVAILABLE WITH SEQLINK

L9 ANSWER 7 OF 33 REGISTRY COPYRIGHT 2003 ACS
RN 485162-38-7 REGISTRY
CN GenBank AAM49872 (9CI) (CA INDEX NAME)
OTHER NAMES:
CN GenBank AAM49872 (Translated from: GenBank AY118503)
SQL 935

SEQ 701 VNDKGVELNH IQINPVWNIT DTYEQGLGTS VSGTVGRIRT STRQFEVMFV
======
HITS AT: 732-737

L9 ANSWER 8 OF 33 REGISTRY COPYRIGHT 2003 ACS
RN 485105-39-3 REGISTRY
CN GenBank AAA46121 (9CI) (CA INDEX NAME)
OTHER NAMES:
CN GenBank AAA46121 (Translated from: GenBank L10087)
SQL 204

SEQ 101 RLRLAFAGVA DGWKGDGYRG HRTGRARHLR LRAATLAASG TVGRERGRTP
== ====
HITS AT: 139-144

RELATED SEQUENCES AVAILABLE WITH SEQLINK

L9 ANSWER 9 OF 33 REGISTRY COPYRIGHT 2003 ACS
RN 483107-64-8 REGISTRY
CN GenBank CAC69156 (9CI) (CA INDEX NAME)
OTHER NAMES:
CN GenBank CAC69156 (Translated from: GenBank AL605486)
SQL 2026

SEQ 451 SGTVGRPECA LHSLSTPEAR TLRGCAKPAP CSDATLASGA VGDAQCSWID
======
HITS AT: 451-456

L9 ANSWER 10 OF 33 REGISTRY COPYRIGHT 2003 ACS
RN 482827-00-9 REGISTRY
CN GenBank AAK59865 (9CI) (CA INDEX NAME)
OTHER NAMES:
CN GenBank AAK59865 (Translated from: GenBank AY037264)
SQL 640

SEQ 351 DIGLLMAGAK ERGELEARVT ALISEVKKSG KVILFIDEVH TLIGSGTVGR

```
                                                                      ======
HITS AT:    395-400

L9   ANSWER 11 OF 33  REGISTRY  COPYRIGHT 2003 ACS
RN   482431-09-4  REGISTRY
CN   GenBank BAA75817 (9CI)  (CA INDEX NAME)
OTHER NAMES:
CN   GenBank BAA75817 (Translated from: GenBank AB018079)
SQL  1249

SEQ    701 VNDKGVELNH IQINPVWNIT DTYEQGLGTS VSGTVGRIRT STRQFEVMFV
                                            ======
HITS AT:    732-737

L9   ANSWER 12 OF 33  REGISTRY  COPYRIGHT 2003 ACS
RN   480853-18-7  REGISTRY
CN   GenBank AAM91802 (9CI)  (CA INDEX NAME)
OTHER NAMES:
CN   GenBank AAM91802 (Translated from: GenBank AY133868)
SQL  945

SEQ    351 DIGLLMAGAK ERGELEARVT ALISEVKKSG KVILFIDEVH TLIGSGTVGR
                                                                ======
HITS AT:    395-400

**RELATED SEQUENCES AVAILABLE WITH SEQLINK**

L9   ANSWER 13 OF 33  REGISTRY  COPYRIGHT 2003 ACS
RN   477113-05-6  REGISTRY
CN   Protein (Mycobacterium leprae clone MLP100854 essential) (9CI)  (CA INDEX
     NAME)
OTHER NAMES:
CN   1849: PN: WO02077183 SEQID: 63849 claimed protein
SQL  911

SEQ    301 EALAAAGERV PEVDEGFDVR GGLLESGTVG RWLAKHADDG RRSGLAIVGT
                                       ===== =
HITS AT:    326-331

**RELATED SEQUENCES AVAILABLE WITH SEQLINK**

REFERENCE   1:  138:1096

L9   ANSWER 14 OF 33  REGISTRY  COPYRIGHT 2003 ACS
RN   462224-65-3  REGISTRY
CN   GenBank AL583921-derived protein GI 13093270 (9CI)  (CA INDEX NAME)
SQL  911

SEQ    301 EALAAAGERV PEVDEGFDVR GGLLESGTVG RWLAKHADDG RRSGLAIVGT
                                       ===== =
HITS AT:    326-331

**RELATED SEQUENCES AVAILABLE WITH SEQLINK**

L9   ANSWER 15 OF 33  REGISTRY  COPYRIGHT 2003 ACS
RN   462224-64-2  REGISTRY
CN   GenBank Z46257-derived protein (9CI)  (CA INDEX NAME)
SQL  911

SEQ    301 EALAAAGERV PEVDEGFDVR GGLLESGTVG RWLAKHADDG RRSGLAIVGT
                                       ===== =
HITS AT:    326-331
```

RELATED SEQUENCES AVAILABLE WITH SEQLINK

L9 ANSWER 16 OF 33 REGISTRY COPYRIGHT 2003 ACS
RN 459753-37-8 REGISTRY
CN GenBank D17582-derived protein GI 497629 (9CI) (CA INDEX NAME)
SQL 945

SEQ 351 DIGLLMAGAK ERGELEARVT ALISEVKKSG KVILFIDEVH TLIGSGTVGR
======
HITS AT: 395-400

RELATED SEQUENCES AVAILABLE WITH SEQLINK

L9 ANSWER 17 OF 33 REGISTRY COPYRIGHT 2003 ACS
RN 457721-02-7 REGISTRY
CN Acyl-CoA dehydrogenase (Mycobacterium tuberculosis strain CDC1551 gene MT3678) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN GenBank AE007169-derived protein GI 13883539
SQL 711

SEQ 101 ATLVVSDPKL RSALASGERF AGVAIDGGVQ VDPKTSTASG TVGRVLGGAP
== ====
HITS AT: 139-144

REFERENCE 1: 137:227411

L9 ANSWER 18 OF 33 REGISTRY COPYRIGHT 2003 ACS
RN 455354-69-5 REGISTRY
CN Transcriptional regulator (Thermosynechococcus elongatus strain BP-1 gene tlr0491) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN GenBank AP005370-derived protein GI 22294212
SQL 219

SEQ 51 TYKVLSGTVG RPQYRYTISE AGREELRQLQ LRQQATHPKG FALELLESVA
===== =
HITS AT: 56-61

REFERENCE 1: 137:211759

L9 ANSWER 19 OF 33 REGISTRY COPYRIGHT 2003 ACS
RN 437143-91-4 REGISTRY
CN ATP-dependent Clp protease ATP-binding subunit ClpD, ERD1 protein precursor (Arabidopsis thaliana clone RAFL09-15-D15 (R09460) gene At5g51070) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN GenBank AY035112-derived protein GI 14334878
SQL 945

SEQ 351 DIGLLMAGAK ERGELEARVT ALISEVKKSG KVILFIDEVH TLIGSGTVGR
======
HITS AT: 395-400

RELATED SEQUENCES AVAILABLE WITH SEQLINK

REFERENCE 1: 137:28849

L9 ANSWER 20 OF 33 REGISTRY COPYRIGHT 2003 ACS
RN 431155-89-4 REGISTRY
CN Protein (Drosophila melanogaster clone WO0171042-SEQID-40731) (9CI) (CA INDEX NAME)
OTHER NAMES:

CN 729: PN: WO0171042 SEQID: 40731 claimed protein
SQL 1249

SEQ 701 VNDKGVELNH IQINPVWNIT DTYEQGLGTS VSGTVGRIRT STRQFEVMFV
======

HITS AT: 732-737

REFERENCE 1: 136:396934

L9 ANSWER 21 OF 33 REGISTRY COPYRIGHT 2003 ACS
RN 401054-29-3 REGISTRY
CN 11: PN: US6350933 SEQID: 11 unclaimed protein (9CI) (CA INDEX NAME)
NTE

type	------ location ------		description
uncommon	Aaa-8	-	-
uncommon	Aaa-22	-	-
uncommon	Aaa-24	-	-
uncommon	Aaa-27	-	-
uncommon	Aaa-33	-	-
uncommon	Aaa-53	-	-
uncommon	Aaa-69	-	-
uncommon	Aaa-82	-	-
uncommon	Aaa-87	-	-
uncommon	Aaa-109	-	-
uncommon	Aaa-112	-	-
uncommon	Aaa-115	-	-
uncommon	Aaa-117	-	-
uncommon	Aaa-125	-	-
uncommon	Aaa-128	-	-
uncommon	Aaa-129	-	-
uncommon	Aaa-209	-	-
uncommon	Aaa-214	-	-
uncommon	Aaa-236	-	-
uncommon	Aaa-238	-	-
uncommon	Aaa-241	-	-
uncommon	Aaa-257	-	-
uncommon	Aaa-263	-	-
uncommon	Aaa-264	-	-
uncommon	Aaa-296	-	-
uncommon	Aaa-317	-	-
uncommon	Aaa-327	-	-
uncommon	Aaa-365	-	-
uncommon	Aaa-385	-	-
uncommon	Aaa-449	-	-
uncommon	Aaa-455	-	-
uncommon	Aaa-784	-	-
uncommon	Aaa-821	-	-
uncommon	Aaa-834	-	-
uncommon	Aaa-848	-	-
uncommon	Aaa-884	-	-
uncommon	Aaa-898	-	-
uncommon	Aaa-904	-	-
uncommon	Aaa-941	-	-
uncommon	Aaa-952	-	-
uncommon	Aaa-955	-	-
uncommon	Aaa-957	-	-
uncommon	Aaa-959	-	-
uncommon	Aaa-960	-	-
uncommon	Aaa-963	-	-
uncommon	Aaa-964	-	-
uncommon	Aaa-969	-	-

uncommon	Aaa-978	-	-
uncommon	Aaa-980	-	-
uncommon	Aaa-987	-	-
uncommon	Aaa-988	-	-
uncommon	Aaa-1003	-	-
uncommon	Aaa-1007	-	-
uncommon	Aaa-1014	-	-
uncommon	Aaa-1019	-	-
uncommon	Aaa-1025	-	-
uncommon	Aaa-1028	-	-
uncommon	Aaa-1035	-	-
uncommon	Aaa-1041	-	-
uncommon	Aaa-1047	-	-
uncommon	Aaa-1060	-	-
uncommon	Aaa-1066	-	-
uncommon	Aaa-1091	-	-
uncommon	Aaa-1092	-	-
uncommon	Aaa-1093	-	-
uncommon	Aaa-1094	-	-
uncommon	Aaa-1097	-	-
uncommon	Aaa-1098	-	-
uncommon	Aaa-1100	-	-
uncommon	Aaa-1101	-	-
uncommon	Aaa-1103	-	-
uncommon	Aaa-1119	-	-
uncommon	Aaa-1126	-	-
uncommon	Aaa-1130	-	-
uncommon	Aaa-1161	-	-
uncommon	Aaa-1165	-	-
uncommon	Aaa-1174	-	-
uncommon	Aaa-1180	-	-
uncommon	Aaa-1191	-	-
uncommon	Aaa-1287	-	-
uncommon	Aaa-1367	-	-

SQL 1402

SEQ 501 VEAKNNLGLS VITYEKPKIE RYEASLVDES GTVGREDDKK KLLEKLLGDK
= =====

HITS AT: 530-535

REFERENCE 1: 136:197195

L9 ANSWER 22 OF 33 REGISTRY COPYRIGHT 2003 ACS
RN 362648-62-2 REGISTRY
CN Protein (Arabidopsis thaliana strain Columbia ecotype gene erd1 dehydration-inducible) (9CI) (CA INDEX NAME)
SQL 945

SEQ 351 DIGLLMAGAK ERGELEARVT ALISEVKKSG KVILFIDEVH TLIGSGTVGR
======

HITS AT: 395-400

RELATED SEQUENCES AVAILABLE WITH SEQLINK

REFERENCE 1: 135:268119

L9 ANSWER 23 OF 33 REGISTRY COPYRIGHT 2003 ACS
RN 290395-99-2 REGISTRY
CN Ferric vibriobactin ABC transporter, permease protein (Vibrio cholerae strain N16961 gene VC0778) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN GenBank AE004163-derived protein GI 9655226

SQL 358

SEQ 151 AGGVMTAIAV YLLSVKSGTV GRYRLVLTGI GVGSVLTALN GLLLVKGSLD
==== ==

HITS AT: 167-172

RELATED SEQUENCES AVAILABLE WITH SEQLINK

REFERENCE 1: 133:218310

L9 ANSWER 24 OF 33 REGISTRY COPYRIGHT 2003 ACS
RN 263502-19-8 REGISTRY
CN Protein (Drosophila melanogaster gene alpha-Man-IIb) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN GenBank AE003710-derived protein GI 7300058
SQL 993

SEQ 701 VNDKGVELNH IQINPVWNIT DTYEQGLGTS VSGTVGRIRT STRQFEVMFV
======

HITS AT: 732-737

REFERENCE 1: 132:289494

L9 ANSWER 25 OF 33 REGISTRY COPYRIGHT 2003 ACS
RN 210885-91-9 REGISTRY
CN Protein (Lactuca sativa gene RG1A) (9CI) (CA INDEX NAME)
NTE

type	location		description
uncommon	Aaa-19	-	-
uncommon	Aaa-83	-	-
uncommon	Aaa-89	-	-
uncommon	Aaa-418	-	-
uncommon	Aaa-455	-	-
uncommon	Aaa-468	-	-
uncommon	Aaa-482	-	-
uncommon	Aaa-518	-	-
uncommon	Aaa-532	-	-
uncommon	Aaa-538	-	-
uncommon	Aaa-575	-	-
uncommon	Aaa-586	-	-
uncommon	Aaa-589	-	-
uncommon	Aaa-591	-	-
uncommon	Aaa-593	-	-
uncommon	Aaa-594	-	-
uncommon	Aaa-597	-	-
uncommon	Aaa-598	-	-
uncommon	Aaa-603	-	-
uncommon	Aaa-612	-	-
uncommon	Aaa-614	-	-
uncommon	Aaa-621	-	-
uncommon	Aaa-622	-	-
uncommon	Aaa-637	-	-
uncommon	Aaa-641	-	-
uncommon	Aaa-648	-	-
uncommon	Aaa-653	-	-
uncommon	Aaa-659	-	-
uncommon	Aaa-662	-	-
uncommon	Aaa-669	-	-
uncommon	Aaa-675	-	-
uncommon	Aaa-681	-	-

uncommon Aaa-694 - -
uncommon Aaa-700 - -

SQL 724

SEQ 151 PKIERYEASL VDESGTVGRE DDKKKLLEKL LGDKDESGSQ NFSIVPIVGM
======

HITS AT: 164-169

REFERENCE 1: 129:145635

L9 ANSWER 26 OF 33 REGISTRY COPYRIGHT 2003 ACS
RN 208861-47-6 REGISTRY
CN Protein FadE34 (Mycobacterium tuberculosis gene fadE34) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN GenBank Z92774-derived protein GI 1877304
SQL 711

SEQ 101 ATLVVSDPKL RSALASGERF AGVAIDGGVQ VDPKTSTASG TVGRVLGGAP
== ====

HITS AT: 139-144

REFERENCE 1: 130:120325

REFERENCE 2: 129:77224

L9 ANSWER 27 OF 33 REGISTRY COPYRIGHT 2003 ACS
RN 198576-75-9 REGISTRY
CN Protein (Methanobacterium thermoautotrophicum strain .DELTA.H gene MTH1013) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN GenBank AE000874-derived protein GI 2622115
SQL 206

SEQ 1 MMSLTPIYID MDGRRVLIVG SGTVGRRRAE RFLKAGAEVA VIGTSEIEGT
======

HITS AT: 21-26

REFERENCE 1: 128:19264

L9 ANSWER 28 OF 33 REGISTRY COPYRIGHT 2003 ACS
RN 196319-84-3 REGISTRY
CN Nucleotidyltransferase, deoxyribonucleate (Mycobacterium leprae clone L247 isoenzyme I gene polI) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN GenBank U00021-derived protein GI 467163
SQL 907

SEQ 301 AAGERVPEVD EGFDVRGGLL ESGTVGRWLA KHADDGRRSG LAIVGTHLPH
======

HITS AT: 322-327

REFERENCE 1: 127:273529

L9 ANSWER 29 OF 33 REGISTRY COPYRIGHT 2003 ACS
RN 169875-44-9 REGISTRY
CN Nucleotidyltransferase, deoxyribonucleate (Mycobacterium leprae clone B961 gene polA) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN DNA polymerase I (Mycobacterium leprae clone B961 gene polA)
CN DNA polymerase I (Mycobacterium leprae strain TN gene polA)
SQL 911

SEQ 301 EALAAAGERV PEVDEGFDVR GGLLESGTVG RWLAKHADDG RRSGLAIVGT
=====

HITS AT: 326-331

RELATED SEQUENCES AVAILABLE WITH SEQLINK

REFERENCE 1: 134:188852

REFERENCE 2: 123:307660

L9 ANSWER 30 OF 33 REGISTRY COPYRIGHT 2003 ACS
RN 158650-69-2 REGISTRY
CN Protein (Saccharomyces cerevisiae gene SPT21 reduced) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN Protein (Saccharomyces cerevisiae gene SPT21)
SQL 758

SEQ 251 NNTNSGTVGR RQTNPMPAPK AVRTQSLPIW NLKPNIANTG FPRNSIAHKI
======

HITS AT: 255-260

RELATED SEQUENCES AVAILABLE WITH SEQLINK

REFERENCE 1: 121:247466

L9 ANSWER 31 OF 33 REGISTRY COPYRIGHT 2003 ACS
RN 152889-46-8 REGISTRY
CN Protein (Arabidopsis thaliana clone ERD1 dehydration-inducible reduced) (9CI) (CA INDEX NAME)
SQL 945

SEQ 351 DIGLLMAGAK ERGELEARVT ALISEVKKSG KVILFIDEVH TLIGSGTVGR
======

HITS AT: 395-400

REFERENCE 1: 120:126392

L9 ANSWER 32 OF 33 REGISTRY COPYRIGHT 2003 ACS
RN 151690-13-0 REGISTRY
CN Protein (Gallid herpesvirus strain 281MI/1 clone pA5 204-amino acid) (9CI) (CA INDEX NAME)
OTHER CA INDEX NAMES:
CN Protein (Marek's disease virus strain 281MI/1 clone pA5 204-amino acid)
SQL 204

SEQ 101 RLRLAFAGVA DGWKGDGYRG HRTGRARHLR LRAATLAASG TVGRERGRTP
== ====

HITS AT: 139-144

RELATED SEQUENCES AVAILABLE WITH SEQLINK

REFERENCE 1: 120:1889

L9 ANSWER 33 OF 33 REGISTRY COPYRIGHT 2003 ACS
RN 107761-90-0 REGISTRY
CN Carboxypeptidase, preproserine (Streptomyces strain R61 clone pDML111 reduced) (9CI) (CA INDEX NAME)
SQL 406

SEQ 1 MVSGTVGRGT ALGAVLLALL AVPAQAGTAA AADLPAPDDT GLQAVLHTAL
======

HITS AT: 3-8

RELATED SEQUENCES AVAILABLE WITH SEQLINK

REFERENCE 1: 123:26572

REFERENCE 2: 106:169914